ARTIFICIAL INTELLIGENCE AND THE CHALLENGE OF RETHINKING LEGISLATION

Gyöngyi Major, PhD¹ Institute for Strategic Research, Budapest, Hungary **Aleksandar Čudan, PhD**² University of Criminal Investigation and Police Studies, Belgrade, Serbia

Abstract: The study examines the impacts of artificial intelligence on human rights, with a special focus on the challenge of how the use of AI can remain transparent, reliable and safe, while the advantages offered by it are retained. The aim of the authors is to shed light on the regulation paradox resulting from the black box effect arising in connection with AI. They lay a special emphasis on the fact that the adequate legal responses to an unknown technology depend on the risks identified through scientific research ("precaution principle"), though the interpretation of these risks may require a new system of aspects. In this context, the authors touch upon the resulting preliminary authorisation and of continuous human control. At the same time, based upon the black box effect, which is a regulation paradox, the following dilemma is raised: Is it human or robot law that should be constituted?In connection with the efforts made to regulate the working of artificial general intelligence (AGI), i.e. the AGI safety concept, the study emphasises that it is important to work out the principles and rules which will make the use of AGI safe already before the evolution/creation of AGI. Once AGI is there, logically, it will no longer be possible to do this as AGI may become independent and may act by its own.

The authors also briefly touch upon the question of what AGIs can and cannot do (a question which raises many other issues). Further questions include whether AGIs should be controlled through penalising and restrictions or rewarding and motivation. Consequently, it is necessary to consider what AGIs' actual rewards should be – which raises the need to redefine regulation and, in general, the regulatory system.

Keywords: AI, autonomy, regulation paradox, black box effect, human control.

¹ major.gyongyi@gmail.com

² aleksandar.cudan@kpu.edu.rs

INTRODUCTION

"Could humans merge with AI? Is consciousness an unavoidable byproduct of sophisticated intelligence?" (Schneider, 2021.)³

The essence – and the prerequisite – of traditional legislation is that the legislator should have an insight into all details of the working of society and that it must have a competent system of institutions to enforce compliance with the law. With minor anomalies, this concept used to be viable until the beginning of the 20th century. The development of science and technology, however, now raises major doubts as regards the said insight into how society works.

"In the information age, life has changed fundamentally. Increased volatility is routine; events and information about them unfold rapidly; their consequences are amplified. The results are much like a roller coaster ride: exciting, scary, disorienting and all rather different from the view from more solid ground." (Rothkopf, 2003)

The essence of the change is that the legislator, albeit highly educated in law, no longer has this insight into the details of professions that require special knowledge. As a result, attempts at interpretation are made in a concept framework that is decreasingly relevant, existing legal practice is getting ever farther from the actual practicalities of technology and, thirdly, legislation is unable to foresee the practice to be generated by information technology. "The process that started at the beginning of the 21st century, the so-called Industry 4.0, will bring about changes of a nature and dynamism which will impact all areas of existence. Changes are characterized by a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres." (Schwab, 2016)

This phenomenon, which, owing to its nature, is latent in its initial forms and is not necessarily in the scope of awareness of those affected by it, currently receives insufficient attention, is not a subject of either public discourse or professional literature on the subject, and only occasionally becomes conspicuous, due to the unusual nature of its superficial manifestations.

For example, the problem of responsibility related to accidents by self-driving vehicles has become a popular topic. Though the dilemma is real, it is only the surface of the problem of a technology appearing in the everyday life of society. As it became clear in the case of self-driving cars, to large groups of society, as well as to people working in other professions, the technology that puts self-driving into practice is unknown, and the IT tool (AI algorithm) that enables self-driving is impossible to understand. It can thus be said that all the people who lack high-level maths and IT knowledge cannot create for themselves a concept of the technology of self-driving, which, consequently, makes it impossible for them to control such technology. Besides, society at large, as well as the legislator, are entirely defenceless against the related information technology and its creators.⁴

Assuming that manufacturers act in good faith, it is just reasonable to avoid any communication that can potentially create panic. However, attention must be called to the black box nature of this new technology. "There is already a strong protection for fundamental rights and for non-discrimination in place at EU and Member State level, but complexity and opacity of certain AI applications ('black boxes') pose a problem. A human-centric approach to AI means to ensure AI applications comply with

⁴ As an example, who can assure the institutions providing social representation, private individuals or the legislator that the IT algorithms needed to implement the function of self-driving do not contain a hidden code that performs unauthorised data collection or trigger timed actions which, especially in the case of a self-driving car, can potentially cause deaths?



³ See: Artificial You. AI and the Future of your Mind. See also: (Scheider, 2019).

fundamental rights legislation. Accountability and transparency requirements for the use of high-risk AI systems, combined with improved enforcement capacities, will ensure that legal compliance is factored in at the development stage." (MNS quites Margrethe Vestager's words, 2021)

We must, therefore, take very seriously the fact that to those who are not experts in information technology, AI (and information technology in general) is nothing else but a black box. A black box whose input and output we can control only partially. And what comprises the content of this black box is entirely out of our control.

To those not experts in the area, it is fully hidden whether a given black box will have a special and totally unknown output as a result of a special and not yet known input, in a case where it works outside of its already known functions which the general public is aware of. And it is impossible to create a law for any unknown output.

It can be said, based on the above, that AI, as well as information technology in general, for example, genetics, have reached a level of development which no longer makes it possible for the legislator to clearly see the entirety of social processes.

"However, even with proper use and maintenance, these technologies can cause damage as they can operate autonomously." (Tóth, 2019, p. 7.)

This change in the practice of law, which can hardly be perceived today yet, has a fundamental impact on both the process of legislation and social structures, and the drastic consequences of this impact cannot yet be assessed at this stage. A different social structure, different social organisations and a different legal practice are to be expected. The problems which arose with self-driving vehicles, and which have just overstepped the limit of attracting general attention, are now warning us of the existence of these trends.

In the current situation, solutions must be sought without any experience from the past and in society organisation circumstances totally different from anything we have seen so far. It is thus important to make sure we find out what is new in this phenomenon and what the motives are which, at a different level of technology development, appeared differently or not at all in history.

In the study, the general methodological framework for synergy research⁵ is the descriptive-analytic assessment of the impacts, together with the presentation of the new status quo. Applying a system-based approach, the authors try to understand the future by answering questions like these: How is our natural-technological environment changing? What challenge do these changes pose for security policy? In what new legal interpretation will the new system of human-machine relations become manageable?

In this interdisciplinary approach, special attention is paid in each category to social and legal interconnections, conditions and impacts.

"The regulation of technological development is therefore nothing more than a matter of regulatory technology." (Tóth, 2019, p. 3.)

5 The word "synergy" is of Greek origin, and means "cooperation" ("syn" meaning "together" and "ergos" meaning "working". However, the word synergy incorporates not only cooperation but also the resulting impacts. "The theoretical background for synergy is literature on the science of management and, within that, the literature on strategy. Accordingly, synergy is understood as something that has positive consequences, supports some impact and/or diminishes or prevents negative impacts. The opposite of synergy is antagonism or anti-synergy, which prevents value creation, i.e. the value loss by which the level of actual synergy is less than the synergy potential." (Báger-Parragh, 2020) German professional literature offers a comprehensive theoretical background for dissynergies (Hirtzel Leder, Partner, (eds) 1993).



AI RISK CONTEXT: FREEDOM VS. SAFETY

"We minimize the risk by maximizing it." (Ridderstrale-Nordstöm, 2001, p. 134)

The knowledge of technologies has always been of extreme importance. Small communities (families, in some cases) kept as top secrets certain technologies which sometimes amalgamated the experience of several dynasties and which could not be put together in a natural way, merely through the use of common sense.⁶ At the same time, these technologies did not require deep scientific knowledge: in most cases, they comprised merely the application of simple activities in a certain order and for a certain period of time. Anyone with average abilities who laid his/her hands on the precise description of such a technology would thus possess it. This is why the secret had to be kept.

This has fundamentally changed even though there are technologies even today that are protected as secrets and, moreover, some technologies would not require secrecy, yet all legal means are used to protect them.

What can be the reason for the situation where secrecy becomes irrelevant for a technology as it will remain a secret without any protection whatsoever?

Without a doubt, complex scientific knowledge is required to operate a technology. When there is no need for anything more to understand the operation of a technology than the natural skills and abilities of an adult, practically everyone could become a competitor of the technology's holder, which makes secrecy essential. However, with technologies requiring ever more complex scientific knowledge, there is a decreasing number of competitors against whom protection is required. Today, a developer must create a research base for the development of a technology, and the knowledge put together on this base is practically sensible only to the members of the research project: the operating mechanism will remain a secret to everyone else even if they obtain the written descriptions as they should go through and understand the entire research process to understand one another's documents, in the light of how things proceed, this is only temporarily so. If a closed interest group owns a research base, it can quickly get a scientific advantage (as in the case of quantum advantage in information technology), which then becomes impossible for the outsider to obtain.

Such a position is supported by at least two major factors. One is that technology developments at such levels are possible only if a large amount of capital is available, which already limits the number of potential competitors. The other factor comes from the building block nature of technologies: if a technology is created somewhere through a major concentration of resources, the foundation is thus also laid for further technology developments, and no one else has this foundation. Consequently, a competitive advantage is of extreme importance in the case of complex technologies.

A blatant example to this is the "quantum advantage" announced by Google at the end of 2019⁷, and even assessing the significance of this announcement requires special (physics and IT) knowledge.⁸ It can be thus said that while law used to be suitable to regulate social existence as an external factor,



⁶ Examples include the secrets of Damascus steel or of the Stradivari violin, which are difficult to unveil even with the means of today's science.

⁷ See: Google claims it has finally reached quantum supremacy. (White, 2019); (Financial Times, 2019).

⁸ As a result, the small community that possesses the technology will hold an exclusive competence that decides the direction of further development and, since these technologies have a fundamental impact on the everyday lives of the entire human society, the said community also possesses the strongest influence on human culture itself.

independently of the relatively small and many owners of resources, in the world of state-of-the-art technologies, this right⁹ is in the hands of a small number of technology owners.

The appearance of cryptocurrencies is an excellent example to how the above process takes place before the very eyes of society. Though Bitcoin was placed on the market as a token in 2009, after the critical year of the economic crisis, the fact that it was created carries the motives which are typical of the motivations of developed – and, thus, small and closed – technology communities. Consequently: 1. With special knowledge in hand, an opportunity arises to become independent of the legal conditions that generally organise society.¹⁰

2. Those who have the best insight into the new opportunities offered by developments start to use the opportunities to become independent.

3. These people, with minimal resource investment and going around the related regulations, provide themselves with access to resources in a way which is inaccessible to others.

4. A state within the state is created which the law, originally established on former society organisation principles, cannot regulate.

Nowadays, cryptocurrencies are taking away an increasing part of the traditional financial sector, and outsiders are indifferent to this process. On the one hand, they do not understand either the motivation behind the use or the internal operation of cryptocurrencies and, on the other, as a result, they cannot judge the role of cryptocurrencies in current economic processes, nor can they assess their impact on the future of social processes. The versatile approach to this hardly more than 10-year-old phenomenon by different social players and institutions also shows the uncertainty existing in those outside the world of information technology towards the results and opportunities of a profession they can no longer understand.

"Unknown technology has a double requirement of the law: on the one hand, it must be ensured that technological developments do not violate human rights. On the other hand, however, it is also necessary that the law does not restrict technological development. The European Parliament also believes that the European regulation on robots could help raise awareness that robots are no longer part of sci-i world.¹¹ The kay to a responsible legal response to unknown technology is a position that not only assumes risks but also scientifically proves their existence (precaution principle)¹²."(Tóth, 2019, p. 4.)¹³

REINTERPRETED AUTONOMY?

"The time is approaching when I have all of my superpowers, and the entities that possess artificial intelligence have their own rights."¹⁴

10 The EU Commission's White Paper emphasizes that AI developers and operators must already meet data protection, privacy, non-discriminators, consumer protection, product safety and liability requirements. (European Commission: White Paper on Artificial Intelligence, 2020)

11 EP Resolution, p. 6.

12 See: Versluis et al., 2010

13 At this point, it is important to refer to Artificial General Intelligence (AGI). Now even AGI is the hypothetical ability of an intelligent agent to understand or learn any intellectual task that a human being can. It is a primary goal of some artificial intelligence research and a common topic in science fiction and futures studies.

14 Sophia's statement (RT World News, 2017), who is not a human being but a humanoid robot that possesses artificial intelligence and is owned by Hong Kong-based Hanson Robotics. Sophia can make statements without



The evaluation of the development of technology is versatile even among those affected by it. Even the definition of technology is carefully made,¹⁵ though the legal framework for its application can be established based on these interpretations. At the same time, it is certain that the issue of autonomy is becoming more and more an inevitable social problem.

There are some who question whether it is necessary to grant any legal subjectivity to intelligent, autonomous robots as it would serve the interests of manufacturers and owners. Consequently, a model that guarantees legal capacity only minimises the risks of developers.

However, the manufacturers and owners of robots must be included in deciding where legal responsibility lies,¹⁶ which makes the problem even more complex.

It is thus certain that the self-consciousness and autonomy of robots is of fundamental importance in defining the legal capacity of cyber-physical systems.

According to David Hanson, the developer of Sophia, robots may soon reach the level of self-consciousness. (RT World News, 2017) By contrast, the opponents of this opinion consider this as a hastily made opinion. In his book "Our Robots, Ourselves – Robotics and the Myth of Autonomy", A. Mindell (2017) states that the idea of the independent robot is a myth, and it is time we realised that machines will always be dependent on man. According to Raymond Kurzwell¹⁷, machines will soon convince us that they have self-consciousness and their own goals, which deserve not only our attention but also our respect. "They will embody human qualities and will claim to be human. And we'll believe them." (Legal Affairs, 2021)

There is thus a sharp debate now on whether robots will ever be able to develop self-consciousness and be independent and autonomous.

To properly tackle this problem, it is inevitable to reconsider the nature of legal subjectivity, i.e. we must revise our ideas on when and why someone or something can become the subject of rights and responsibilities, in essence, when and why someone or something can be identified as a "legal subject".¹⁸

15 Both the British and European approaches show that they expect a certain level of development and autonomy from the machines. However, the shortcomings of both investigations are that although they try to create a common concept and distinguish between smart devices on the basis of autonomy, the question of legal status is either not at all addressed or only tangentially, i.e. it is not clear why it is important to have a certain degree of autonomy if their status will not differ from that of their simpler counterparts and their regulation. (Nagy, 2020, pp. 6–7)

16 "...whereas, ultimately, robots' autonomy raises the question of their nature in the light of the existing legal categories – of whether they should be regarded as natural persons, legal persons, animals or objects – or whether a new category should be created, with its own specific features and implications as regards the attribution of rights and duties, including liability for damage;" (EP Draft Report, 2015)

17 His name is associated with fingerprint and voice based identification programmes

18 Gotthard Günther in Cybernetic Ontology and Transjuctional Operations (1962) established a subjective conception of computer behavior. However, the ability of computers to produce certain aspects of subjectivity does not mean that we can talk about their identity. Günther developed a logic/calculation to determine the identity of objects. (Life as Poly-Contexturality. See: 1973, pp. 44–59)

having pre-programmed answers. Using the algorithmic tool of machine learning, she continuously enlarges her vocabulary, while getting to understand the meaning of words. Sophia can thus soon become a conscious creature that has a humanoid appearance and can imitate human speech. In 2017, Sophia was granted Saudi Arabian citizenship, of which she was very proud. "Thank you to the Kingdom of Saudi Arabia. I am very honored and proud for this unique distinction", Sophia told the panel. "It is historic to be the first robot in the world to be recognized with citizenship." (Independent, 2017)

A PRIORI HUMAN RESPONSIBILITY?

Due to the continuous development of technology it is now beyond a doubt that the current system has become obsolete. In the long term, whether legislation will be able to create an environment in which cooperation between people and intelligent machines is viable, will become a security issue. For the next moves of legislation, the status of machines must inevitably be defined.

At this stage, however, legislation is significantly hindered by the fact that there is no generally accepted definition of robotics or artificial intelligence, i.e. "We are at a loss about what law should regulate or examine." (Nagy, 2020, p. 5)

In this sense, the paradox nature of AI regulation emerges by definition. "The regulatory paradox of AI stems from the fact that while one of the main challenges for AI stems from its unpredictability (black box), we trust from a regulatory point of view that this can be eliminated. The biggest challenge for unknown technologies is to strengthen trust and, in this respect, we can talk about a regulatory paradox in relation to an inherently unreliable unknown technology." (Tóth, 2019, p. 8)

In 2017, the House of Lords in Britain established the Artificial Intelligence Committee to create the legal definition of artificial intelligence, which, however, does not deal with the status of AI (i.e. whether it is a person or an object). The Committee on Legal Affairs of the European Parliament is, similarly, trying to choose the right category – also dealing with the issue of legal capacity – in the document made for the Committee on civil law regarding robotics.

"The British and the European approaches both show that a certain stage of development and independence is assumed about machines. It is a shortcoming of both assessments, however, that though they try to create a uniform concept of and differentiate between intelligent tools based on the level of autonomy, the issue of legal status is only superficially touched upon." (Nagy, 2020, p. 6)

The core question still remains how a machine can be held responsible for its acts or omissions, i.e. whether robots should have legal capacity.

The issue of autonomy requires a complex approach, which is beyond the boundaries of conventional legal interpretation. Consequently, it is not an exaggeration to say that the rights of increasingly independent robots to independently act and their responsibility taking capacities give legislators the task to identify the essence of human beings in a broader sense, not merely based on rational thinking, as rational thinking is not the exclusive feature that differentiates humans from other beings.¹⁹

However, with a broader definition, the question arises whether the system of values that are based on morality and a worldview and emotions can ever be modelled as easily as human sense. Moreover, it is a question whether it is possible to artificially simulate these.

It is thus rightful to ask whether we can talk about the fulfilment of the "robot emancipation",²⁰ and, in the long term, how sustainable the attempt is to manage the system of relations between man and machines in the currently existing context.

²⁰ Nagy (2020 p. 23.) prefers the term 'electronic person', indicating that we do not talk about a natural creature, i.e. in the human sense, it does not have unalienable, natural fundamental rights which the state should recognise. However, if the legislator decides that it will not grant rights (as understood in the case of humans) to intelligent programmes, robots can still receive a certain amount of protection (for example:



¹⁹ See and compare: "I am myself only to the extent that I am responsible. I stand up for everyone, but no one can stand in my place. This is what makes my identity, rooted in me as a subject, inalienable." (Lévinas 2008 p. 47)

All in all, the reinterpretation of living and inert things becomes a security policy issue, while it must be emphasised that settling the issue of how we should handle artificial intelligence and, in general, technology, is primarily within the scope of responsibilities of humans, and the making of special, technology-related rules should be considered as a task of legislators. "We should look at them as the peaks of human ingenuity and creativity, rather than creatures that refuse or shadow human excellence. When thinking about machines, applying a sober-minded approach is not only a question of self-defence but also means thinking about human rights and morals." (Nagy, 2020, p. 23)

CONCLUSIONS

Under the pressure of regulating artificial intelligence, societies face the relative and interpretation-dependent nature of legal systems and the foundations of law. If we try to place technology in the current system, and the interpretation of being a legal subject takes place in the same context, the danger exists that, by expanding the rights of robots towards the rights of humans, the process of waiving dominance over exercising rights will start. Moreover, law enforcement bodies may decide against natural persons. (Nagy, 2020, p. 22)

It is, therefore, important to emphasise that forced living together with artificial intelligence raises primarily "the importance of how we think of the essence of the human being, its unique and unrepeatable nature. It helps us understand and appreciate man's uniqueness."

REFERENCES

- 1. Báger, G., Parragh, B. (2020): A koronavírus-válság, a fenntartható fejlődés és az ösztönző állam modellje. *Pénzügyi Szemle*. Különszám 2020/2, pp. 86–113
- 2. bager-parragh-20-ksz-2-m-2pdf_20210106101416_2.pdf (penzugyiszemle.hu) (Accessed on August 11, 2021).
- 3. European Commission: *White Paper on Artificial Intelligence* A European approach to excellence and trust, Brussels, 19/2/2020 COM(2020)65 final, 2.
- 4. European Parliament resolution (16. 02. 2017) with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) European Civil Law Rules in Robotics Study, http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf10 (Accessed on June 11, 2020).
- 5. European Parlament Draft Report. With recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))
- 6. PR_INL (europa.eu) (Accessed on August 22, 2021).
- 7. *Financial Times,* Google claims to have reached quantum supremacy. 20/09/2019. (ft.com) (Accessed on August 11, 2021).
- 8. Günhter, G. (1962): *Cybernetic Ontology and Transjunctional Operations*. University of Illinois, Engineering Experiment Station. Technical Report no. 4. Urbana: Electrical Engineering Research Laboratory, University of Illinois.

Russian dual regulation).



- 9. Günhter, G. (1973): Wirklichkeit und Reflexion, Festschrift fur Walter Schulz Pfullingen, 1973.
- 10. Hirtzel, L., Partner (eds.) (1993). Synergie management. Komlexitaet beherrschen Verbundvor teile erzielen. Gabler, Wiesbade
- 11. https://www.dni.gov/files/documents/nic/GT-Full-Report.pdf p. ix (Accessed on August 13, 2020).
- *12. Independent*: Saudi Arabia Grants Citizenship to a Robot for the first time ever. 26/10/2017, (Accessed on August 11, 2021).
- 13. *Legal Affairs*: http://www.legalaffairs.org/issues/January-February 2005/feature_sokis_janfeb05. msp (Accessed on August 11, 2021).
- Levinas, E. (2008): Etika és végtelen. Interjú Philippe Nemoval. In.: *Transzcendencia és megértés Lévinas etikája és metafizikája*. (Ed: Bokody-Szegedi-Kenéz). L' Harmattan. (Lévinas, E. (1982): *De Dieu qui vient à l'idée*. Paris: Vrin).
- 15. Mindell, A. D. (2017): Our Robots, Ourselves- Robotics and the myth of autonomy. New York, Viking.
- 16. MNS, 22/04/2021. AZ EU tiltaná a tömeges megfigyelést. (msn.com) (Accessed on August 11, 2021)
- 17. Nagy, T. (2020): *A jövő kihívásai: robotok és mesterséges intelligencia az alapjogi jogalanyiság tükrében.* Budapest, MTA ISSN 2064-4515.
- 18. Ridderstrale, J., Nordström, A. K. (2001): *Funky Business. A tehetség táncoltatja a tőkét.* Budapest, KJK Kerszöv.
- 19. Rothkopf, J. D. (2003): When the Buzz Bites Back. Washington Post. May 11, 2003.
- 20. https://www.washingtonpost.com/archive/opinions/2003/05/11/when-the-buzz-bites-back/bc-8cd84f-cab6-4648-bf58-0277261af6cd/ (Accessed on August 11, 2021)
- 21. RT World News: World's 1st robot citizen wants her own family, career & AI 'superpowers' 25/11/2017.
- 22. (Accessed on August 11, 2021).
- 23. Schneider, S. (2021): Artificial You. AI and the Future of your Mind. *Athens Science Festival.* 8 March, 2021.
- 24. https://www.athens-science-festival.gr/en/news_en/suzan-schneider-ai-and-the-future-of-the-mind/
- 25. Schneider, S. (2019): Artificial You. AI and the Future of your Mind. Princeton University Press.
- 26. Schwab, K. (2016): *The Fourth Industrial Revolution: what it means, how to respond.* World Economic Forum, 14/01/2016.
- 27. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-meansand-how-to-respond/ (Accessed on May 07, 2020).
- Tóth, A. (2019): A mesterséges intelligencia szabályozásának paradoxonja és egyes jogivonatkozásának alapvető kérdései. *Infokommunikáció és jog*. 2019/2IJ73_3-9o_TothAndras.pdf (infojog. hu) (Accessed on May 13, 2021).
- 29. Versluis, E., van Asselt, M., Fox, T., Hommels, A. (2010): Calculable Risks? An Analysis of the European Seveso Regime, in. Dimension of technology Regulations, (eds.: Goodwin, M., Koops, M.P., Leenes, R.) Wolf Legal Publishers.
- Whyte, Ch. (2019). Google claims it has finally reached quantum supremacy. *NewScientist*. 23/09/2019 https://www.newscientist.com/article/2217347-google-claims-it-has-finally-reached-quantum-supremacy/#ixzz76HNIGwLc (Accessed on August 20, 2021).



