ROBOTS WITH BIOLOGICAL BRAINS: AUTONOMY AND LIABILITY OF A SEMI-ARTIFICIAL LIFE FORM

Alžběta Krausová,* Hananel Hazan**

Abstract: *Hybrid IT systems with biological brains (hybrots) enhance our understanding of brain functioning. However, given their specific form of existence and their ability to act autonomously to a certain degree, they raise questions regarding attributing liability in case they cause damage. The aim of this paper is to suggest a scheme for attributing liability to these systems while taking into account liability issues related to artificial intelligence. Firstly, the paper describes the technology of hybrid systems, its history, state of the art as well as its technological constraints. Next, general social and ethical aspects of the technology are briefly described. Finally, the paper describes the main problems related to liability of hybrid systems and sets out criteria for determining liability for damages caused by hybrots.*

Keywords: *artificial intelligence, autonomy, brain-computer interface, bio-robotic organism, cybernetic organism, ethics, hybrid robot, hybrot, liability, neuromorphic robot, responsibility gap, robot, semi-artificial life, semi-organic species*

INTRODUCTION

We as a human kind have always been fascinated by ourselves, namely by our existence and the origin of our extraordinary capabilities. Out of them an ability to think and solve problems has been up to now considered as the most valuable faculty of a human. Therefore, many thinkers of all centuries have been trying to find out what makes them think, how they think, what limits them in thinking, and in which ways they can put their thoughts into action.

With regard to this problem, the 20th century brought a scientific breakthrough – a possibility to study thinking through testing hypotheses on artificial subjects: neuromorphic and brain-based robots. These robots are characterized as *"physical devices whose control system has been modelled after some aspect of brain processing*".[1] William Grey Walter's turtles, that were originally constructed in 1948-1949, represent one of the first robots of this kind. These small autonomous devices exhibit quite a complex behavior. Although their electrical brain contains only two "neurons", they are capable of phototaxis. Their creator called them a new in-organic species – Machina speculatrix.[2]

Since then, scientific knowledge has advanced rapidly and currently we are able to go beyond simulating brain functioning; we are able to interface living brains with artificial bodies and even with new brains created out of living neurons. These robots belong among so called hybrid systems and can be considered as a new semi-organic species. They are also sometimes called as cybernetic or bio-robotic organisms.

* Mgr. Alžběta Krausová, LL.M. Institute of Law and Technology at the Faculty of Law, Masaryk University, Brno, Czech Republic. Institute of State and Law of the Czech Academy of Sciences, Prague, Czech Republic

** Hananel Hazan, Ph.D. Network Biology Research Laboratories, Technion, Israel

1 KIRCHMAR, J. L., WAGATSUMA, H. (Eds.). *Neuromorphic and Brain-Based Robots*. Cambridge: Cambridge University Press, 2011, p. 4.

2 PICKERING, A. *The Cybernetic Brain. Sketches of Another Future.* Chicago: The University of Chicago Press, 2010, p. 43.

These hybrid systems are capable of autonomous behavior and together with Artificial Intelligence (AI) bring a completely new element into a social interaction. Due to their relative independence and, therefore, reduced control over them on the side of their creators and users, the society is facing a problem of how to guarantee safety and stability that has traditionally been secured through the institute of legal responsibility. The aim of this paper is to assess applicability of the existing liability schemes on hybrots.

The remainder of the paper is organized as follows: firstly, hybrid systems with biological brains, their history, functioning, current state-of-the-art, and technological constraints that we are now facing shall be introduced. Next, the relationship between AI and hybrots will be depicted. Subsequent chapter will focus on general questions, namely reasons for constructing hybrots, relevant risks posed by them, and a way of approaching the problem. Then, the question of social impact of autonomy of hybrid systems and AI will be analyzed from the perspective of law, focusing on liability for damage resulting from autonomous behavior of these systems. Finally, the paper shall conclude with policy recommendations based on findings from previous chapters.

1. HYBRID IT SYSTEMS WITH BIOLOGICAL BRAINS

A computer science defines a hybrid IT system as a system which combines and processes data from a continuous input (an input that continuously varies within predefined boundaries) with data from a discrete input (an input that is not continuous).[3] In hybrid IT systems with biological brains the discrete input is based on electrical signals representing neural activity of a monitored brain.

There are basically two types of hybrid systems that utilize brain power. The first type is a brain-computer interface. Brain-computer interfaces (BCIs) are *"communication systems that do not depend on the brain's normal output pathways of peripheral nerves and muscles. In these systems, users explicitly manipulate their brain activity instead of using motor movements to produce signals that can be used to control computers or communication devices"*.[4] These systems use fully functioning brains embedded in biological bodies as sources of a discrete input. The second type is a so called "hybrot". This hybrid system utilizes electrical activity of a brain that has been artificially created out of living biological neurons and preserved in vitro as a source of a discrete input.

The term hybrot originates from the words "hybrid robot". The term was coined for the first time by a research group led by Steve M. Potter from the Georgia Institute of Technology in 2003. Hybrot was referring to *"[living neuron's] cultures interfaced to physical robots"*[5] and distinguished a new type of a system from systems that use the term "animat".

[3] BRANICKY, M. S. Introduction to Hybrid Systems. In HRISTU-VARSAKELIS, D. and W. S. LEVINE (Eds.). *Handbook of Networked and Embedded Control Systems.* Boston: Birkhäuser, 2005, p. 91.

[4] TAN, D. S., NIJHOLT, A. (Eds). *Brain-Computer Interfaces. Applying our Minds to Human-Computer Interaction.* London: Springer, 2010, p. 4.

[5] BAKKUM, D. J., SHKOLNIK, A. C., BEN-ARY, G., GAMBLEN, P., DeMARSE, T. B., POTTER., S. M. Removing Some 'A' From AI: Embodied Cultured Networks. In IIDA, F., PFEIFER, R., STEELS, L., KUNIYOSHI, Y. (Eds.). *Embodied Artificial Intelligence. International Seminar, Dagstuhl Castle, Germany, July 7-11, 2003, Revised Selected Papers.* Berlin: Springer, 2004, p. 133.

An animat was defined as *"a computer simulated or robotic animal behaving in an environment"*.[6] The first hybrot was constructed by the above mentioned research group already in 2001 but at that time it was still called *"neurally controlled animat"*.[7] Hybrot is currently defined simply as a *"robot controlled by living neurons"*.[8]

The purpose of constructing a hybrot was to learn about how neural tissues work in real-time, how they learn and react, and how computer science and namely the field of artificial intelligence could benefit from the new findings.

Development of hybrots was preceded by interesting ideas and other types of robots. Research was primarily inspired by the work of Valentino Braitenberg.[9] In 1984, Valentino Braitenberg proposed *"an exercise in fictional science, or science fiction"*[10] in which he contemplated how simple machines equipped just with few sensors and directed only by several simple rules would operate in a natural environment. He concluded that operation of such machines may seem as intelligent and pointed out that properties of these machines originate in properties of animals' brains. These machines were later called Braitenberg's vehicles.

In 1990s, a second generation of Braitenberg's vehicles, so called BEAM robotics[11] was invented. BEAM robots have simple analogue brains and are oriented only to survival (i.e. they can move in an unknown environment). Nevertheless, their behavior becomes quite complex.[12]

Exploring "intelligent behavior" was brought forth a decade later with hybrid systems integrating living neural tissue into a robotic system. Before the first hybrot was constructed, the neuroscience had already had a substantial body of knowledge on how a neural tissue behaves when isolated from a body. Therefore, in 2000 Reger and his colleagues managed to create a hybrid system in which a portion of a brain of a sea lamprey *Petromyzon marinus* was interfaced with an artificial body – a small mobile Khepera robot.[13] This system fulfils the above mentioned definition of a hybrot (which did not yet exist then) since the robot's behavior was directed by the lamprey's brain that was stimulated through electrodes based on input from light sensors.

---

6 Ibid., p. 133.

7 DeMARSE, T. B., WAGENAAR, D. A., BLAU, A. W., POTTER, S. M. The Neurally Controlled Animat: Biological Brains Acting with Simulated Bodies. *Auton Robots*. 2001, Vol. 11, No. 3, p. 305.

8 POTTER, S. M. et al. Hybrots: Hybrids of Living Neurons and Robots for Studying Neural Computation. *Proceedings of Brain Inspired Cognitive Systems*. 2004, [2017-06-01]. Available at: <http://s3.amazonaws.com/academia.edu.documents/30019486/1094.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1497906963&Signature=h62RxeY%2B5Y5XEO6DRmBr1Fz2UCA%3D&response-content-disposition=inline%3B%20filename%3DHybrots_Hybrids_of_Living_Neurons_and_Ro.pdf>.

9 MAROM, S., MEIR, R., BRAUN, E. et al. On the Precarious Path of Reverse Neuro-engineering. *Frontiers in Computational Neuroscience*. 2009, Vol. 3, p. 1.; REGER, B. D., FLEMING, K. M., SANGUINETI, V. et al. Connecting Brains to Robots: An Artificial Body for Studying the Computational Properties of Neural Tissues. *Artificial Life*. 2000, Vol. 6, No. 4, p. 323.

10 BRAITENBERG, V. *Vehicles. Experiments in Synthetic Psychology*. Cambridge: The MIT Press, 1984, p. 1.

11 The abbreviation BEAM is standing for Biology, Electronics, Aesthetics and Mechanics.

12 HASSLACHER, B., TILDEN, M. W. Living Machines. In: *Scholarbotics.net* [online]. 14. 6. 1994 [2017-01-20]. Available at: <http://www.solarbotics.net/library/pdflib/pdf/living_machines.pdf>.

13 REGER, B. D., FLEMING, K. M., SANGUINETI, V. et al. Connecting Brains to Robots: An Artificial Body for Studying the Computational Properties of Neural Tissues. *Artificial Life*. 2000, Vol. 6, No. 4, p. 308.

Later, various hybrots based on different methods were constructed for instance by DeMarse et al. in 2001,[14] Bakkum et al. in 2003,[15] Martinoia et al. in 2004,[16] Marom et al. in 2009,[17] Mussa-Ivaldi et al. in 2010,[18] or Warwick et al. in 2010.[19] In principle, these hybrots consisted of three main parts: a biological brain, hardware, and software.

The core of all these hybrots was a biological brain that was artificially created out of rats neurons. To construct such a "brain", a part of an original rat's brain counting approximately tens of thousands of brain cells is firstly disassociated with help of enzymes and mechanic tools and later *"cultured on a Petri dish with […] electrodes embedded in the substrate, a multi-electrode array […]. The neurons in these cultures spontaneously branch out […] Even left to themselves without external input other than nutrients (cell culture media), they re-establish connections with their neighbors and begin communicating electrically and chemically within days, demonstrating an inherent goal to network*".[20] This part of a hybrot is called "wetware" and, after being subject to the described process, can be thought of as tabula rasa, i.e. a clean state in which no connections are yet made so the brain functioning can be studied from the very beginning. However, a hybrot also needs hardware and software components in order to function in our world.

The multi-electrode array (MEA) together with a stimulation system are the key hardware components which ensure bi-directional communication between the cultured neural tissue and its artificial body. MEA detects and records neurons' activity, namely their firing. The recorded signals (spikes) are analyzed by specific algorithms and used as commands for an artificial body. On the other hand, input from outside sensors is translated into electrical signals and delivered as feedback back to the neural tissue again through the MEA. Other hardware components may include a mobile robotic body typically equipped with sensors or a static IT system with a simulated virtual environment that provides feedback to the wetware.

Relevant software forms an integral part of a hybrot. Its purpose is in general to acquire and manage signals from the wetware. The crucial components are a decoding scheme and a coding scheme. The decoding scheme comprises a clustering algorithm that learns

---

[14] DeMARSE, T. B., WAGENAAR, D. A., BLAU, A. W., POTTER, S. M. The Neurally Controlled Animat: Biological Brains Acting with Simulated Bodies. *Auton Robots*. 2001, Vol. 11, No. 3, pp. 305–310.

[15] BAKKUM, D. J., SHKOLNIK, A. C., BEN-ARY, G., GAMBLEN, P., DeMARSE, T. B., POTTER., S. M. Removing Some 'A' From AI: Embodied Cultured Networks. In IIDA, F., PFEIFER, R., STEELS, L., KUNIYOSHI, Y. (Eds.). *Embodied Artificial Intelligence. International Seminar, Dagstuhl Castle, Germany, July 7–11, 2003, Revised Selected Papers*. Berlin: Springer, 2004.

[16] MARTINOIA, S., SANGUINETTI, V., COZZI, L. et al. Towards an Embodied In Vitro Electrophysiology: the NeuroBIT Project. *Neurocomputing*. 2004, Vol. 58–60, pp. 1065–1072.

[17] MAROM, S., MEIR, R., BRAUN, E. et al. On the Precarious Path of Reverse Neuro-engineering. *Frontiers in Computational Neuroscience*. 2009, Vol. 3, pp. 1–4.

[18] MUSSA-IVALDI, F. A., ALFORD, S. T., CHIAPPALONE, M. et al. New Perspectives on the Dialogue between Brains and Machines. *Frontiers in Neuroscience*. 2010, Vol. 4, No. 1, pp. 44–52.

[19] WARWICK, K., XYDAS, D., NASUTO, S. J. et al. Controlling a Mobile Robot with a Biological Brain. *Defence Science Journal*. 2010, Vol. 60, No. 1, pp. 5–14.

[20] BAKKUM, D. J., SHKOLNIK, A. C., BEN-ARY, G., GAMBLEN, P., DeMARSE, T. B., POTTER., S. M. Removing Some 'A' From AI: Embodied Cultured Networks. In IIDA, F., PFEIFER, R., STEELS, L., KUNIYOSHI, Y. (Eds.). *Embodied Artificial Intelligence. International Seminar, Dagstuhl Castle, Germany, July 7–11, 2003, Revised Selected Papers*. Berlin: Springer, 2004, p. 132.

to recognize patterns in the neural activity according to frequency of spikes at each location. The coding scheme, on the other hand, processes information from sensory inputs, codes them and sends signals to corresponding areas in the wetware. In some cases learning rules are also implemented that either reward or punish the wetware for a certain behavior.[21]

Over the past 15 years many research groups have been developing hybrots with the aim to study neural functioning. Steve M. Potter's group has devised three setups illustrating possibilities how hybrots could be utilized in the future: a) a system in which living neurons controlled movement of a virtual body in a virtual environment; b) a system in which living neurons controlled a mobile robot; and c) a system in which living neurons controlled a drawing robotic arm that was geographically detached and located 18.105 kilometers far from the place where the "brain" was based.[22]

What concerns the state-of-the-art and achievements in the field, we can say that hybrots are currently mainly capable of movement, they are able to follow or avoid a light source, to escape obstacles, or to navigate a flight. Progress has been achieved by some researchers in inducing plastic changes in cultured neural tissues that promote quicker learning.[23]

The current research, however, faces constraints that hinder the development of hybrots. Although we try to study brain functioning in real-time by monitoring neuronal activity, we are not able *"to map the path from stimuli to action through the brain"*.[24] So far, the essence of information conveyed by neural electrical activity is not understood. This limitation of knowledge has impact on the way in which software for hybrots is programmed.

Since scientists need to determine interpretation of neuronal activity, they design an algorithm that has two principal functions. Firstly, the algorithm cleans out noise[25] from recorded neural activity. This step is quite problematic because, despite the fact that cleaning out noise is necessary, it may eliminate potentially important information. Secondly, the algorithm determines which patterns of activity shall influence hybrots' behavior and in which way. By using such algorithm researchers influence the system's behavior although the neuronal tissue is otherwise free to act autonomously, i.e. create connections and fire nerve impulses.

---

[21] TESSADORI, J., BISIO, M., MARTINOIA, S., CHIAPPALONE, M. Modular Neuronal Assemblies Embodied in a Closed-loop Environment: Toward Future Integration of Brains and Machines. In POTTER, S. M., EL HADY, A., FETZ, E. E. (Eds.). Closing the Loop Around Neural Systems. In: *Frontiers* [online]. 2014 [2017-01-20]. Available at: <http://www.frontiersin.org/books/Closing_the_Loop_Around_Neural_Systems/386>. See p. 86.

[22] BAKKUM, D. J., SHKOLNIK, A. C., BEN-ARY, G., GAMBLEN, P., DeMARSE, T. B., POTTER., S. M. Removing Some 'A' From AI: Embodied Cultured Networks. In IIDA, F., PFEIFER, R., STEELS, L., KUNIYOSHI, Y. (Eds.). *Embodied Artificial Intelligence. International Seminar, Dagstuhl Castle, Germany, July 7–11, 2003, Revised Selected Papers.* Berlin: Springer, 2004, p. 134 et seq.

[23] TESSADORI, J., BISIO, M., MARTINOIA, S., CHIAPPALONE, M. Modular Neuronal Assemblies Embodied in a Closed-loop Environment: Toward Future Integration of Brains and Machines. In POTTER, S. M., EL HADY, A., FETZ, E. E. (Eds.). Closing the Loop Around Neural Systems. In: *Frontiers* [online]. 2014 [2017-01-20]. Available at: <http://www.frontiersin.org/books/Closing_the_Loop_Around_Neural_Systems/386>. See p. 86.

[24] MAROM, S., MEIR, R., BRAUN, E. et al. On the Precarious Path of Reverse Neuro-engineering. *Frontiers in Computational Neuroscience*. 2009, Vol. 3, p. 1.

[25] The noise in this sense means activity of brain cells that does not fulfill certain criteria defined by designers and is, therefore, filtered out and not reflected in the behavior.

Future research presumes improvement of hybrots' abilities through increasing number of neurons in a tissue, constructing three dimensional brains, broadening autonomy by provision of more effective power sources to mobile robots, installing more kinds of sensors and fusion of perceived information, or development of methods enabling greater control over learning and memory.[26]

What concerns future applications of hybrots, there are not yet any specific proposals. In general, scientists presume utilizing neuronal robustness and biological computers in development of intelligent machines. Moreover, some visionaries speculate that the ongoing research can be one day used as groundwork for achieving immortality by placing a living human brain into an artificial body that transcends a fragile biological nature of a human body.

2. NATURAL AND ARTIFICIAL INTELLIGENCE

Originally, it was a biological brain whose capabilities and structure inspired the field of AI to study and simulate neural functioning. Over the past nearly 60 years AI has achieved outstanding results and has been widely used for all kinds of purposes.[27] However, it seems that AI is returning to its roots as it is now used to analyze and learn from real neural functioning.

Although the technology of hybrots is currently far from being mature and compared to AI quite inefficient, it might have a very promising future. So far we know that a biological brain is *"a supremely versatile, efficient, capable, robust and intelligent machine"*[28] and that Natural Intelligence (NI) as such is in its complexity superior to AI. Therefore, scientists suggested that AI should get inspired by NI, namely by following neuroscientific research and implementing specific features of biological brains into its architecture since NI differs from AI in many characteristics.[29]

What concerns the future, it seems as the most logical estimation that one day both NI and AI shall assist each other more and more up to the degree that they might merge.

3. SOCIETY AND DEVELOPMENT OF HYBROTS

The very concept of a hybrot is controversial. It poses many ethical questions, a number of which were identified by Warwick. As the first problem Warwick points out a trauma of the biological tissue caused by taking away part of its original identity (a biological body) and providing the tissue with a completely new environment. This could be either a completely disembodied environment with no sensory input or interfacing the tissue with

[26] WARWICK, K., XYDAS, D., NASUTO, S. J. et al. Controlling a Mobile Robot with a Biological Brain. *Defence Science Journal*. 2010, Vol. 60, No. 1, p. 11.

[27] LUNGARELLA, M., IIDA, F., BONGARD, J., PFEIFER, R. (Eds.). *50 Years of Artificial Intelligence. Essays Dedicated to the 50th Anniversary of Artificial Intelligence.* Berlin: Springer, 2007.

[28] POTTER, S. M. What Can AI Get From Neuroscience? In LUNGARELLA, M., IIDA, F., BONGARD, J., PFEIFER, R. (Eds.). *50 Years of Artificial Intelligence. Essays Dedicated to the 50th Anniversary of Artificial Intelligence.* Berlin: Springer, 2007, p. 174.

[29] Ibid.

a system providing the tissue some feedback that should be, according to scientists, meaningful to the tissue. Since neurons fire even when deprived of sensory input, Warwick poses questions that are up to now impossible to answer, such as *"what does its body mean to the culture?"*.[30] Warwick adds that some people hope that we can probably resolve such problem one day by using human neurons. However, this seems to be even a more provoking idea, especially because some scientists envision the technology to enable us to preserve our minds in robotic bodies.

Experimenting with different kinds of neurons raises specific questions, such as whether in the future we shall have various semi-organic "species", whether we need to distinguish neurons from various animals, whether we should avoid or promote species discrimination, or whether there is some limitation on number of neurons that we should use in order to limit or, on the other hand, to possibly increase capabilities of hybrots.

Despite the problems mentioned, there are many reasons why we should want to develop hybrots. The main argument is that hybrots provide us with a unique ability to perform specific experiments helping us in studying life. According to Warwick, *"studying such neuronal systems can help us to understand biological neural structures in general and it is to be hoped that it may lead to basic insights into problems such as Alzheimer's and Parkinson's disease"*.[31] This opportunity should not be thrown away just because there are risks involved.

Moreover, people have always been motivated by the desire to understand the life and the world and to exercise control over it. Realistically, the research will not stop. The society will need to deal with upcoming questions raised by hybrots not only from a philosophical, ethical, or medical perspective, but also from a legal perspective.

The role of law in society is among others to seamlessly integrate new findings and technologies into everyday life so we all can benefit from them. Despite there are many questions raised by hybrots, this paper shall focus only on a question of liability for their behavior and analyze remaining issues in later papers. The reason for this is a fact that the first and most pressing question our society will need to deal with shall regard practical matters. Such as in case of many technologies in the past, development of hybrots challenges a fragile balance between the social interest to ensure safety in the society and an interest to promote potentially risky research. A potential solution can be found in refining the concept of liability.

4. LIABILITY FOR AUTONOMOUS BEHAVIOR OF HYBRID SYSTEMS

Hybrid systems represent a specific interconnection of three basic components – a biological brain, hardware that is manipulated by a brain activity, and software that enables communication between the brain and the hardware. From a legal perspective, it is necessary to distinguish the following two forms of hybrid systems: brain-computer interfaces

[30] WARWICK, K. Implications and Consequences of Robots with Biological Brains. *Ethics of Informational Technology*. 2001, Vol. 12, No. 3, p. 228.
[31] Ibid., p. 224.

and hybrots. Brain-computer interfaces are currently used mainly for monitoring human brains that manipulate the hardware or for research of brain activity in animals. BCIs are using subjects that are recognized by law and, therefore, liability rules derive from their legal status. Hybrots, on the other hand, represent a completely new entity that cannot be simply considered as a thing because a part of this entity is a living tissue. At the same time one cannot consider this entity as an animal because it functions in a completely different manner. Moreover, hybrots are equipped with AI that not only enables learning of a hybrot, but it can also interfere significantly with her brain activity.

Artificial intelligence itself has been considered to challenge the very concept of liability and will, therefore, be examined as well.[32] Liability is a concept closely related to the concept of responsibility. Responsibility refers to specific qualities of a person, namely to consideration and care for one's own actions. Being responsible can be defined on several levels and can refer to being *"answerable or accountable, as for something within one's power, control or management"*,[33] *"having a capacity for moral decisions and therefore accountable; capable of rational thought or action"*,[34] or being *"able to discharge obligations or pay debts"*.[35] Being liable, on the other hand, means being legally responsible. Responsibility is in this sense enforceable by law in the form of a legal obligation.

Liability can refer both to a legal obligation of a person to behave in a certain manner in order not to cause damage as well as to a legal obligation to face negative consequences in case a person either fails to fulfill her legal obligation and causes damage (for instance breach of contract or negligence), or is in charge of circumstances under which damage is caused even though she have not caused it herself (strict liability).[36] In general, the purpose of liability is to define rules for undesirable situations involving damage. These rules determine who shall remedy such situations.

From a societal point of view, liability serves several crucial functions. First of all, liability rules help to safeguard a peaceful state in the society through securing predictability of results of possible judicial proceedings. Liability also serves a restorative function (a legal obligation to restore the original state before the damage was caused), a compensatory function (a legal obligation to reimburse the caused damage), a repressive function (a punishment for the person who caused damage), and a preventive function (liability rules motivate people not to cause damage by implying the obligation to restore or compensate the damage as well as to receive punishment).[37]

[32] There is a number of publications that point out specificities of AI and its implications on liability. See for example FRANK, S. J. Tort Adjudication and the Emergence of Artificial Intelligence Software. *Suffolk University Law Review*. 1987, Vol. XXI, pp. 623–667; COLE, G. S. Tort Liability for Artificial Intelligence and Expert Systems. *Computer Law Journal*. 1990, Vol. 10, pp. 127–231; GERSTNER, M. E. Liability Issues with Artificial Intelligence Software. *Santa Clara Law Review*. 1993, Vol. 33, pp. 239–269; VLADECK, D. C. Machines Without Principals: Liability Rules and Artificial Intelligence. *Washington Law Review*. 2014, Vol. 89, pp. 117–150.

[33] *Webster's Encyclopedic Unabridged Dictionary of the English Language*. New York : Random House Value Publishing, Inc., 1996, p. 1641.

[34] Ibid.

[35] Ibid.

[36] HARVÁNEK, J. a kol. *Právní teorie*. [*Legal Theory*]. Plzeň: Aleš Čeněk, 2013, p. 358–359.

[37] Ibid., p. 360.

In principle, only natural and legal persons can be held liable for their actions. It is due to the fact that law is a social construct created for humans. Only humans are currently able to understand a legal system and use it properly.[38] The capacity to be held liable depends to a high degree on comprehension of how the world and the society function, while the capacity to comprehend depends on mental faculties of a human. Humans are in this sense autonomous beings with the ability to understand social relationships and with the ability to decide about their actions. They can be held liable for constructing and using things and instruments that are in principle in their control.

However, development of highly intelligent information processing AI applications having the freedom to make own decisions and the ability to act autonomously up to a certain degree initiated discussions about problems with causality and attributing legally relevant behavior to the originators of a particular AI software. With regard to a lack of control of developers over the AI software, so called *"responsibility gap"* has been identified.[39]

This responsibility gap refers to a situation in which a machine operates by itself and makes her own decisions without any further interference of a human. If there are any negative consequences of such operation, generally either a manufacturer or an operator of the machine is held liable. However, if they have no control over decision made by a machine, holding them liable seems unjust. Manufacturers are, therefore, undoubtedly motivated to advocate a limitation of their liability. Making them liable for something that is beyond their control may result in unwillingness to continue research and consequently hinder AI development. The same can be said about hybrots since they are also autonomous to a certain degree.

As the current liability framework gradually becomes unsatisfactory, it is necessary to analyze possible solutions in the light of the new developments. At present, there are several models of attributing liability for damages caused by AI that can be considered.

A first model implies no liability for damages caused by AI. As it was mentioned earlier, at a certain moment developers stop to exercise full control over decisions made by AI. The behavior of AI becomes unpredictable and, theoretically, it could be compared for instance to force majeure or an act of God. However, this model is highly inconvenient as to the values that the current modern society holds onto. In the end, such approach would destabilize legal relationships in the society as it would not provide any solution for undesirable situations. Moreover, AI systems are to a certain degree controllable by various subjects who create them and later influence them over the course of their lifetime. The same is also true for hybrots.

Another model that should be, therefore, considered is a model of strict liability or a model of liability without fault. In this case law holds liable explicitly determined subjects that have a specific relationship with AI/hybrots which have caused damage. This model is currently preferred as it ensures that all the previously described social functions of liability are fulfilled.

[38] The capacity of legal persons to enter legal relationships and to be held liable for their actions is derived from the capacity of natural persons who establish and govern such legal person.

[39] MATTHIAS, A. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*. 2004, Vol. 6, No. 3, p. 177.

However, recently the European society started to consider whether, depending on a level of intelligence achieved, AI could become a special kind of a person who could bear liability for her own actions such as natural or legal persons. In the past such reflections were considered as pure science fiction that has no reasonable justification in our society. Nevertheless, in early 2017, the European Parliament adopted a resolution with recommendations on Civil Law Rules on Robotics.[40] In this resolution the European Parliament describes the rapid technological development in the field of robotics and AI and points out namely the progress in achieving autonomy and cognitive abilities of robots. The EP raises question whether it would not be useful to create a new legal status *"with its own specific features and implications"*[41] as over time the model of strict liability may not be sufficient. Currently robots with AI are considered as products and their producers can be liable for damages only if these are caused as a result of a defect in a robot. This causal link will be impossible to prove. Moreover, such defect cannot be proved in hybrots as their biological brain cannot be efficiently shaped. Therefore, creating a specific legal status for hybrots is even more justifiable than in the case of AI.

When considering which model of liability is appropriate for both AI and hybrots, it is necessary to bear in mind several criteria and specificities. Artificial intelligence as such is programmable and its creators are free to determine the level of its autonomy by limiting its knowledge, freedom to overwrite certain rules etc. Therefore, if AI can never break some rules, it has no meaning to setup a liability scheme that would somehow punish the system. The purpose of punishment is to influence and modify behavior of a person in the future to prevent possible misconduct. Moreover, in general AI does not suffer from loss of self-control such as humans who can fail in preventing and not causing damage for instance because of their imbalanced mental state.

This, however, is not valid for hybrots, because their actions are directed by the activity of a neural tissue. A manufacturer of such hybrot has little control over how the neural tissue will behave. A manufacturer of a hybrot can influence mostly the AI that processes data gained by monitoring the neural tissue. There is some evidence on a possibility to shape functioning of a neural tissue by stimulating it with electric signals. However, this stimulation is far from guaranteeing absolute control over the tissue. Anyway, in case of hybrots, it seems that a punishment and reward scheme could work in order to fulfill the preventive function of liability.

Apart from this, there are three main potential risk factors when dealing with AI and hybrots: context of usage, maturity of technology, and an initial setting of an organic brain. From the point of view of setting up a liability scheme one needs to be aware of the fact that different contexts and environments have different requirements on safety. In a laboratory environment hazard is unavoidable since research cannot move forward without trying the unknown. For use by wide public safe solutions are necessary. Maturity of AI and hybrots en-

[40] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)). In: *European Parliament* [online]. 7. 3. 2017 [2017-05-20]. Available at: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//EN>.

[41] Ibid.

sures relative reliability and depends on time devoted to learning that a particular technology undergoes. Immature technologies are again unavoidable in labs and unacceptable for wide use. With regard to biological brains, brains with structures formed earlier may not be controllable as effectively as new and clean artificially created brains.

Having these factors in mind, the model of strict liability seems suitable for distinguishing different stages in research and development of the technology. This model can differentiate rules for research environments from rules for producers of systems for use by general public. Researchers could be protected from higher risks in laboratory environments by being held liable only for ignoring or neglecting precautions but not for unfortunate results. On the other hand, producers of commercial systems for general public should be held liable also for some unfortunate results in order to be motivated to implement highest possible security measures into these systems.

Moreover, in case the technology gets mature enough, it is valid to start considering a model of setting up a specific status of an "electronic person" or in case of hybrots a specific status of a "hybrid person". The reason for distinguishing a hybrid person from electronic person is the biological core of such a person. This biological core not only enables this person to act autonomously and, therefore, unpredictably, but it also enables this person to come up with autonomy of own goals, intentions and desires that are not probable to be achieved by a purely electronic person.

Finally, the concept of a hybrid person is justifiable because biological brains of hybrots call for a specific additional type of liability: a liability for harming hybrots. With regard to the biological core of a hybrot, we must presume that this entity can suffer. For instance, the more sensory inputs hybrots shall have, the more sensitive they will become. For the same humane reasons, we protect animals from torture, we need to protect hybrots. Many questions, however, remain as to the hybrid person's capabilities, level of intelligence and species discrimination with regard to the origin of neural tissues, their size and organization.

CONCLUSION

The previous sections have illustrated that autonomous behavior of AI and hybrots challenges the current liability schemes. One of the possible legal solutions is to keep and refine the model in which liability is attributed to more subjects while respecting specific risks posed by various environment-, purpose- and developmental stage related conditions. This model should hold liable all subjects dealing with AI and hybrots for the periods during which these subjects exercise their limited control over the technologies to avoid situations without obligation to restore order.

At the same time, it is now possible to consider setting out a specific status of an electronic person that would supplement the current statuses of a natural and legal persons. With regard to the specificities of hybrots and their biological core, in the future we should consider also a status of a hybrid person who, in fact, operates in a different manner and as opposed to an electronic person equipped only with AI. A hybrid person can have own goals and desires and her behavior is modifiable within the traditional terms of reward and punishment. Without being able to have own legal status such hybrid person would not be able to possibly compensate any damage caused by her and the manufacturer or the owner of this hybrot might not be held liable.