COMPUTER LAW & SECURITY REVIEW 31 (2015) 376–389

Available online at www.sciencedirect.com

ScienceDirect

www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

Liability for damages caused by artificial intelligence

Paulius Čerka [a], *Jurgita Grigienė* [a,*], *Gintarė Sirbikytė* [b]

[a] *Čerka and Partners Advocates Bureau, Vytautas Magnus University, Faculty of Law, Department of Private Law, Kaunas, Lithuania*

[b] *Čerka and Partners Advocates Bureau, Kaunas, Lithuania*

Keywords:
Artificial intelligence
Liability for damages
Legal regulation
AI-as-Tool
Risks by AI
Respondeat (respondent) superior
Vicarious liability
Strict liability

A B S T R A C T

The emerging discipline of Artificial Intelligence (AI) has changed attitudes towards the intellect, which was long considered to be a feature exclusively belonging to biological beings, i.e. homo sapiens. In 1956, when the concept of Artificial Intelligence emerged, discussions began about whether the intellect may be more than an inherent feature of a biological being, i.e. whether it can be artificially created. AI can be defined on the basis of the factor of a thinking human being and in terms of a rational behavior: (i) systems that think and act like a human being; (ii) systems that think and act rationally. These factors demonstrate that AI is different from conventional computer algorithms. These are systems that are able to train themselves (store their personal experience). This unique feature enables AI to act differently in the same situations, depending on the actions previously performed.

The ability to accumulate experience and learn from it, as well as the ability to act independently and make individual decisions, creates preconditions for damage. Factors leading to the occurrence of damage identified in the article confirm that the operation of AI is based on the pursuit of goals. This means that with its actions AI may cause damage for one reason or another; and thus issues of compensation will have to be addressed in accordance with the existing legal provisions. The main issue is that neither national nor international law recognizes AI as a subject of law, which means that AI cannot be held personally liable for the damage it causes. In view of the foregoing, a question naturally arises: who is responsible for the damage caused by the actions of Artificial Intelligence?

In the absence of direct legal regulation of AI, we can apply article 12 of United Nations Convention on the Use of Electronic Communications in International Contracts, which states that a person (whether a natural person or a legal entity) on whose behalf a computer was programmed should ultimately be responsible for any message generated by the machine. Such an interpretation complies with a general rule that the principal of a tool is responsible for the results obtained by the use of that tool since the tool has no independent volition of its own. So the concept of AI-as-Tool arises in the context of AI liability issues, which means that in some cases vicarious and strict liability is applicable for AI actions.

* *Corresponding author*. Vytautas Magnus University Faculty of Law, Department of Private Law (Lithuania), Ožeskienės 18, LT-44246 Kaunas, Lithuania.
E-mail address: jurgita.grigiene@cerka.lt (J. Grigienė).
http://dx.doi.org/10.1016/j.clsr.2015.03.008
0267-3649/

1. Introduction

Rapid technological change and development has led to an era of complex Artificial Intelligence (AI) technology and applications.[1]In addition to the positive effects of AI, such as increased production and indirectly lowered costs in factories and production lines, reduced potential of errors and increased efficiency, replacement of human labor in dangerous situations, etc.; but along with progress in technology come negative outcomes as well. Operation of AI is a new phenomenon that sometimes is not sufficiently understood. AI is different from conventional computer algorithms in that it is able to train itself on the basis of its accumulated experience. This unique feature enables AI to act differently in the same situations, depending on the actions previously performed. Therefore, in most cases the efficiency and potential of AI is rather unclear. Nevertheless, AI is increasingly being used in everyday life. Sometimes, people enjoy the benefits of AI without even realizing it (smart, self-training programs, such as Siri, bots, etc.).

The prevalence of AI in society means that the more people use AI, the greater the likelihood of various violations of law. Accordingly, AI development and its ever-growing practical use require changes in legal regulation, such as the need to restructure the legal system. If Artificial Intelligence turns out as planned, i.e. a thinking human-like robot with feelings and emotions, then the laws would need to be altered to encompass the roles of robots in society.[2] It means that lawmakers must review the existing legal framework and adapt it to the changing needs of society.

The increasing use of AI in the field of public activities leads to some challenges in the area of legislation. Daily use of IT, including AI, by the network society is different in its principles of operation from institutional, often bureaucratic, behavior. The operating system of IT and its integral part AI is spontaneous, constantly evolving and changing. For this reason, legislation governing this field should be: (i) universal in order to be effective, regardless of changes in information technology, or (ii) constantly amended in order to be effective, regardless of changes in information technology. However, the second option, i.e. constant change in legislation depending on changes in information technology, may be difficult to implement due to static and consistent nature of operation of institutions.

In 2012, the European Commission initiated a RoboLaw Project with the main objective of investigating the ways in which emerging technologies in the field of bio-robotics (including AI) bear on the national and European legal systems, challenging traditional legal categories and qualifications, posing risks to fundamental rights and freedoms that have to be considered, and more generally demanding a regulatory ground on which they can be developed and eventually launched. The most important outcome of the RoboLaw Project appeared on the 22 September, 2014. It consists of a final report containing “Guidelines on Regulating Robotics”, addressed to the European Commission, in order to establish a solid legal framework for the development of robotic technologies in Europe.

However the problem of AI legal regulation has to be solved not only in Europe. It is obvious that while assessing the effect of accelerating globalization processes, the problem of AI cannot be limited by territoriality and its highlighting of different legal traditions practices. The lack of legal regulation in the field of AI is a problem for the global citizenship – all of network society, including civil law and common law countries as well. This problematic extends beyond national borders, which mean that it is not a problem of individual country or continent. This is a worldwide significance problem. For this reason we need not only the regional unification act of AI law, but the global one as well.

While the operation of AI is not regulated by specific legislation, we have to deal with the issue of liability for damages caused by the actions of AI. Legal norms provide that the damages caused by unlawful actions of others must be compensated. Legal norms provide that the damage is to be compensated by the offender or a person who is responsible for the actions of the offender. In view of these legal regulations and the fact that AI is not the subject of law yet, a question arises: who is required to assume liability and compensate damages caused by the actions of AI? This question, moreover, describes *the aim of the research*, i.e. to find out who is to be held liable for the damage caused by the Artificial Intelligence. *The object of the research* is the liability of Artificial Intelligence for the damages caused by its actions. *The methods of the research* are: information collection, systematizing, generalizing, valuation, comparison, analysis of scientific literature and legal acts, synthesis and deduction.

Some of the main questions to answer are: is AI capable of causing damage (is it possible that AI may be hazardous and may cause damage)? Can AI be held liable for its actions? What is the legal regulation to be employed in identifying the cases of compensation of damages caused by AI? Topics about AI in general and about the liability of AI for its illegal actions have been studied by some authors with differing conclusions.[3]

[1] Abdul Ahad Siddiqi, ‘Implications of using Artificial Intelligence Technology in Modern Warfare’ [2012] Communications and Information Technology (ICCIT), 2012 International Conference, 30; <http://www.taibahu.edu.sa/iccit/allICCITpapers/pdf/p30-siddiqi.pdf>.

[2] Richard C. Sherman, ‘The Surge of Artificial Intelligence: Time To Re-examine Ourselves. Implications: Why Create Artificial Intelligence?’, (1998) <http://www.units.muohio.edu/psybersite/cyberspace/aisurge/implications.shtml>’accessed 16.11.13.

[3] For example: David C. Vladeck, (2014); Ugo Pagallo *The Law of Robots* (2013); Ray Kurzweil, *The Singularity Is Near* (2005); Stuart Russell & Peter Norvig, *Artificial Intelligence Modern approach* (2009); Rymantas Tadas Toločka, *Regulated Mechanisms*(2008); Michael Aikenhead, *Legal Analogical Reasoning – the Interplay Between Legal Theory and Artificial Intelligence* (1997); Nils J. Nilsson, *The Quest for Artificial Intelligence. A history of ideas and achievements* (2010); Miglė Laukytė, *Artificial and Autonomous: A Person?* (2012); Marshal S. Willick *Constitutional Law and Artificial Intelligence: The Potential Legal Recognition of Computers as “Persons”* (985); Kathleen Mykytyn & Peter P. Mykytyn, & Jr. Craig W. Slinkman, *Expert Systems: A Question of Liability* (1990); Luke Muehlhauser & Anna Salamon, *Intelligence Explosion: Evidence and Import* (2012); Curtis E.A. Karnow, *Liability For Distributed Artificial Intelligences* (1996); Maruerite E. Gerstner, *Liability Issues with Artificial Intelligence Software* (1993); Erica Palmerini, *The Interplay Between Law and Technology, or the RoboLaw Project in Context* (2012); Geoffrey Samuel, *The Challenge of Artificial Intelligence: Can Roman Law Help Us Discover Whether Law is a System of Rules?* (2006); Stephen M. Omohundro, *The Basic AI Drives* (2007.

The article consists of three main parts. The first part introduces the concept of the operation of Artificial Intelligence software, discusses its meaning and significance, and elaborates on its operating principles. Moreover, it reflects on the practical aspect, discusses risks and potential cases of damages caused by AI as well as identifies factors that precondition the occurrence of damages. The first part of the article, moreover, presents examples of the existing AI systems and the risks posted by them. The second part briefly presents an overview of the problematic of AI legal regulation. And the third part analyses the concept of AI-as-Tool in the context of AI liability. In this section the articles deals with issues of respondeat superior liability theory, vicarious and strict liability application for AI actions, in terms of the fact that general principle on article 12 of United Nations Convention on the Use of Electronic Communications in International Contracts states that a person (whether a natural person or a legal entity) on whose behalf a computer was programmed should ultimately be responsible for any message generated by the machine.

It is important to note that information technologies, including activities of AI, are highly dynamic and constantly changing. Such dynamism and uniqueness causes most of the problems, because scientific literature that deals with the relevant issues becomes outdated rather quickly. It should be noted that although the popularity and the scope of use of AI is growing, there are no specific laws that define and regulate their operation. Thus, regardless of the specificity of AI, the general principles of law and legislation governing the operation of IT should be employed in the meantime.

2. The concept of artificial intelligence

2.1. *What is artificial intelligence?*

Artificial Intelligence (AI) is a new discipline[4] that has no universally accepted definition[5] and formally emerged in 1956, when the concept of Artificial Intelligence was mentioned for the first time.[6] The emergence and rapid development of personal computers led to the creation and research of AI. Until now, the prevailing view was that intellect is an exclusive feature of a biological being, i.e. homo sapiens.[7] This belief began to change due to constant improvement of computer systems. Intellect, i.e. the ability to know, understand, and think, may not only be innate (natural), but also artificially created. AI is a broad area that includes such subfields as: (i) natural language systems; (ii) machine learning; (iii) simulation of senses; (iv) neural networks; (v) computer games; (vi) expert systems; and (vii) robotics.[8]

The definition of Artificial Intelligence found in articles by various authors[9] states that AI is any artificially created intelligence, i.e. a software system that simulates human thinking on a computer or other devices: e.g. home management systems integrated into household appliances; robots; autonomous cars; unmanned aerial vehicles, etc. Something that some decades ago was only science fiction, such as software systems dealing with various issues and replicating thinking processes of humans by means of hardware and other technology, has become the reality of science today.[10]

Stuart J. Russell and Peter Norvig identify two main trends in AI definitions. Fig. 1 One is associated with the processes of thinking and motivation, and the other is linked to behavior.[11] According to Russell and Norvig, AI can be defined as systems that think, and systems that act and rationalize like humans.[12]

In view of the above definitions, it is clear that AI is different from conventional computer algorithms. The development of Artificial Intelligence is aimed at making it self-training (the ability to accumulate personal experience) or machine learning. This unique feature enables AI to act differently in the same situations, depending on the actions previously performed. This is very similar to human experience. Cognitive modeling and rational thinking techniques give more flexibility and allow for the creation of programs that can "understand," i.e. that have the traits of a reasonable person (brain activity processes). Hayes[14] argues that "in their complexity and operating principles, certain software programs imitate certain processes of the human brain," the functioning of which is based on an artificial neural network.

Based on the aforementioned, it is safe to say that the aim of AI developers is to teach a computer to master the human intellect and think as well as act rationally and intelligently later on. This aim can be achieved through the following fundamentals which Ph. D. K. M. Tadiou defines as (i) Problem solving by search; (ii) Knowledge representation; (iii) Machine

[4] Stuart Russell, Peter Norving, *Artificial Intelligence: A Modern Approach* (NJ Prentice Hall 1995) 3.

[5] Rymantas Tadas Toločka, *Regulated Mechanisms* (Technologija 2008) 30.

[6] Vytautas Čyras, *Artificial Intelligence* (E-book 2008) 5, <http://www.mif.vu.lt/~cyras/AI/konspektas-dirbtinis-intelektas.pdf> accessed 25.11.13.

[7] Russell, Norving (n 4) 3.

[8] Ibid; John McCarthy, 'What is artificial intelligence?'(2007 Stanford University, Computer Science Department) 15. <http://innovation.it.uts.edu.au/projectjmc/articles/whatisai/whatisai.pdf>accessed 25.08.13.

[9] Ibid; Gabriel Hallevy, 'The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control' (2010) 4 Akron Intellectual Property Journal 171, 175–178, <https://www.uakron.edu/dotAsset/1139176.pdf> accessed 28.08.13.

[10] Maruerite E. Gerstner, 'Liability Issues with Artificial Intelligence Software' (1993) 33(1) Santa Clara L. Rev., 239 <http://digitalcommons.law.scu.edu/lawreview/vol33/iss1/7> accessed 25.09.13, Citing *Cariad Hayes*, 'Artificial Intelligence: The Future's Getting Closer', (1988) AM. LAW, 115.

[11] Stuart Russell, Peter Norving, *Artificial Intelligence: A Modern Approach*.(3rd edn, NJ Prentice Hall 2009) 1–2.

[12] Ibid 2.

[13] Russell, Norving (n 11) 2.

[14] Hayes (n 9), Cited in *Gerstner* (n 9).

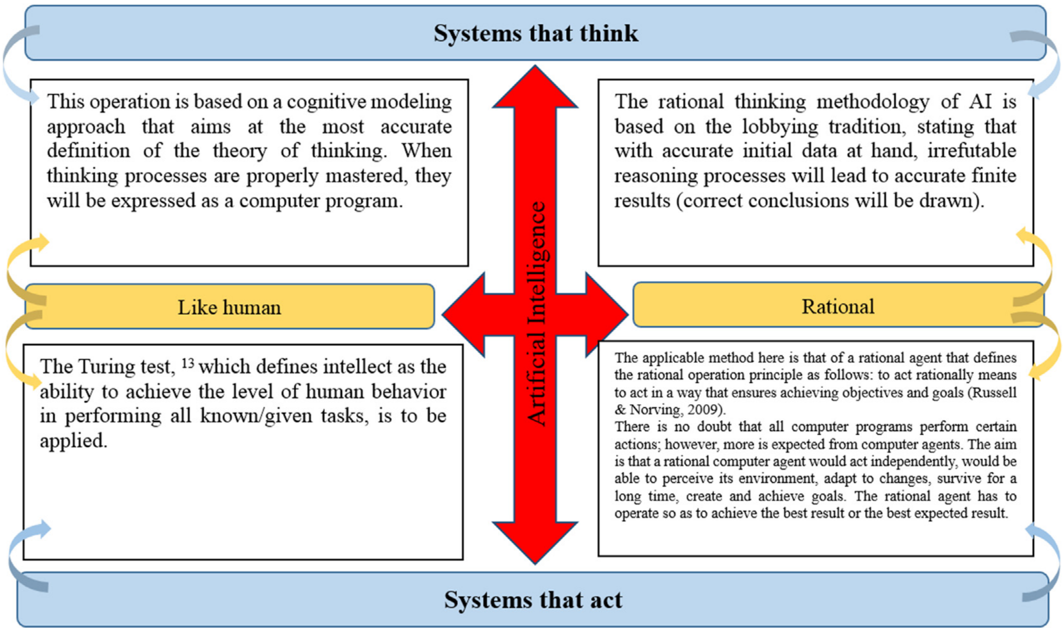

Fig. 1 – Definitions of AI by S. J. Russell & P. Norving[13].

learning; (iv) Natural language processing; and (v) Artificial neural networks.[15]

Aspects of search underlie much of artificial intelligence. When a problem is given to AI to solve, AI developers do not give a specific algorithm for how to solve the problem, but only a description that lets the AI recognize a solution. So the task for AI is to search for a solution.[16] The idea of search is straightforward. Fig. 2 Tadiou it calls a *tree* with its roots (initial state), nodes (states), branches (one or several legal moves) and leaves (terminal state) which are used to represent the set of all possible legal moves and can be explored breadth first or depth first.[17]

The figure above illustrates that "AI constructs a set of potential partial solutions to a problem that can be checked to see if they truly are solutions or if they could lead to solutions. Search proceeds by repeatedly selecting a partial solution, stopping if it is a path to a goal, and otherwise extending it by one more arc in all possible ways."[19]

Knowledge representation is central to AI research and development.[20] The diversity of subfields of the knowledge representation range is unified by the central problem of encoding human knowledge in such a way that the knowledge can be used.[21] AI has to solve problems which require extensive knowledge about the world. This is why artificial intelligence knowledge must be acquired, and represented in special language like first order logic[22] and coded in order to make it possible for machines to manipulate.[23] So knowledge representation is the area of AI concerned with how knowledge can be represented symbolically and manipulated in an automated way by reasoning programs.[24]

Ronald J. Brachman and Hector J. Levesque paraphrase the words of philosopher Brian Smith by saying that the "Knowledge Representation (KR) Hypothesis implies that we will want to construct knowledge-based systems for which the intentional stance is grounded by design in symbolic representations."[25] The authors point out that the KR hypothesis supports two main and general properties of which the first is to understand knowledge-based systems as standing for propositions and the second is that the system is designed to behave the way that it does because of these symbolic representations.[26]

Another fundamental part of AI is the *Machine learning* process, the aim of which is to give AI the ability to learn new facts from data with no explicit program and to adapt things it

[15] Koné Mamadou Tadiou, 'Introduction to Artificial Intelligence',<http://futurehumanevolution.com/artificial-intelligence-future-human-evolution/introduction-to-artificial-intelligence> accessed 19May 2014.

[16] David Poole, Alan Mackworth, *Artificial Intelligence: Foundations of Computational Agents* (Cambridge University Press, 2010) (Web version) <http://artint.info/html/ArtInt_47.html> accessed 23.05.14.

[17] Tadiou (n 16).

[18] Bill Wilson, 'Problem Solving and Search in AI' (2002) <http://www.cse.unsw.edu.au/~billw/cs9414/notes/mandc/mandc.html> accessed 22.03.14.

[19] Poole, Mackworth (n 17).

[20] Russell, Norving (n 11) 234–321; Randall Davis, Howard Shrobe, Peter Szolovits, 'What is a Knowledge Representation?' (1993) *AI Magazine*, 14 (1):17–33 <http://groups.csail.mit.edu/medg/ftp/psz/k-rep.html>accessed 23.05.14.

[21] Christopher A. Welty, 'An Integrated Representation for Software Development and Discovery', Disertation (1996) <http://www.cs.vassar.edu/~weltyc/papers/phd/HTML/dissertation-1.html> accessed 23.05.14.

[22] Russell, Norving (n 11) 285.

[23] Tadiou (n 16).

[24] Ronald J. Brachman, Hector J. Levesque, *Knowledge Representation and Reasoning* (2003) (Web version) <http://rair.cogsci.rpi.edu/pai/library/brachmanbook7-17-03.pdf> accessed 23.05.14.

[25] Ibid 6.

[26] Ibid 5.

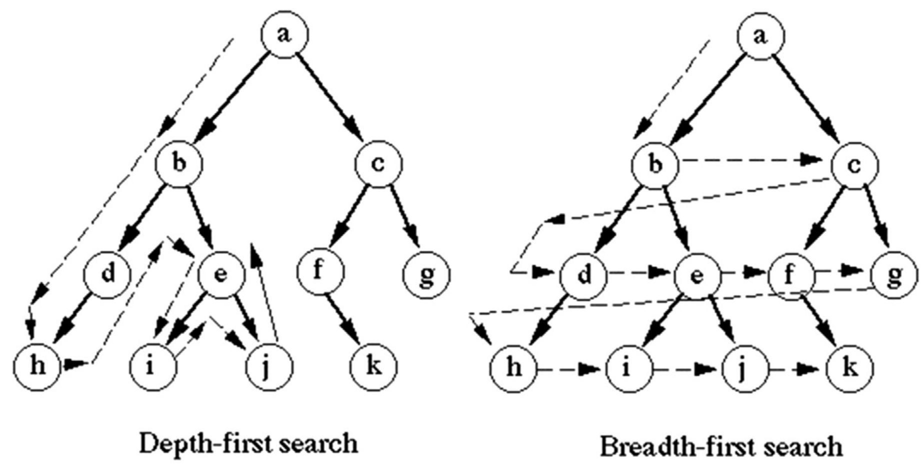

Fig. 2 – Two alternative methods of searching a tree by Bill Wilson[18].

has learned to a new situation. This definition of Machine learning was authored by Arthur Samuel in 1959.[27] Flying-robot—the 2002 project of Chalmers University of Technology scientists—is a good example of the machine learning process. A broader discussion of this project is presented below.

Natural Language Processing is a subfield of the component of AI that allows a computer program to understand human speech as it is spoken.[28] One of the challenges inherent in natural language processing is human–computer interaction through the understanding, processing and elicitation of human language to teach computer to understand the way humans learn and use language.[29]

In addition to what was mentioned above, another fundamental part of AI is *Artificial neural networks* (ANN). The Computing Dictionary defines ANN as a network of many simple processors—"neurons" (nodes), each of which possibly has a local memory. Fig. 3 The nodes are connected by unidirectional communication channels ("links"), which carry numeric data. The nodes operate only on their local data and on the inputs they receive via the links.[30] But it is not just a complex system; ANN is a complex *adaptive* system, meaning it can change its internal structure based on the information flowing through it.[31]

ANN is like the natural human brain with its biological neurons and synapses, the goal of which is to reproduce the computing power of the human brain.[33] A network of many nodes can exhibit incredibly rich and intelligent behaviors such as ability to *learn*.[34]

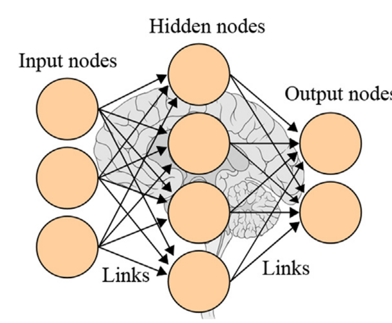

Fig. 3 – The basic structure of an Artificial Neural Network, by PhD K. M. Tadiou[32].

The development of all AI subfields (fundamentals) allows AI to cover many areas: from the analysis of understanding and logical reasoning to advice on tax matters, aircraft piloting, driverless cars, public safety (police), computer configuration, the game of chess, proving mathematical theorems, writing poetry and books, drawing paintings and diagnosing disease, language decipherment and signature recognition.[35] Developing research in the field of AI suggests that the potential of AI will only increase,[36] which naturally raises the question of the liability of AI and compensation of damages, if and when AI for one reason or another causes damages with its actions, i.e. by accidentally or intentionally making errors, etc.

[27] Cited in Phil Simon, Too Big to Ignore: The Business Case for Big Data (Wiley 2013) 89.
[28] Margaret Rouse, 'natural language processing', (2011)<http://searchcontentmanagement.techtarget.com/definition/natural-language-processing-NLP> accessed 12.06.14.
[29] Tadiou (n 16).
[30] Denis Howe, artificial neural network. Dictionary.com, *The Free On-line Dictionary of Computing*, <http://dictionary.reference.com/browse/artificial neural network > accessed 12.06.14.
[31] Daniel Shiffman, 'The nature of code', (2012) <http://natureofcode.com/book/chapter-10-neural-networks/> accessed 12.06.14.
[32] Tadiou (n 16), *Adapted from two pictures on Wikipedia*.
[33] Tadiou (n 16).
[34] Shiffman (n 32).
[35] Gerstner (n 9).
[36] Nils J. Nilsson, *The Quest for Artificial Intelligence. A history of ideas and achievements* (Web version 2010) 589, <http://ai.stanford.edu/~nilsson/QAI/qai.pdf>accessed 12.12.13.

2.2. Practical aspect: capabilities of AI: independent decision-making

The likelihood of damages caused by the actions of AI is real. Already there have been cases where the liability for certain actions of a person is passed on to Artificial Intelligence.[37] Although currently Artificial Intelligence systems possessing only primitive (weak) properties are used in practice, the manifestations of independent decision-making by AI as well as the unlimited development and learning potential of AI have been noted.[38] In order to properly demonstrate the significance of AI, this paper presents examples of AI used in practice.

As early as 2002, scientists at Chalmers University of Technology (Sweden) successfully trained a robot in flying basics. A winged robot was made that could move its wings up/down and back/forward thereby rotating about its axis. The robot was attached to two poles and programmed to achieve maximum lift, but it did not have an algorithm instructing it how to do this. The winged robot simply used its evolutionary self-training algorithm. Initially, the robot was performing random movements, the results of which were assessed by means of a feedback; unsuccessful movements as well as their combinations were rejected, while the successful ones were recorded and used to improve the flapping technique. It took the robot 3 h to learn to flap its wings and achieve the lift.[39] While it learned how to flap its wings, the robot showed it could palter (i.e. "cheat"). Having detected objects accidentally placed under its wings, it used these objects to lift its wings. After a couple of hours, the robot flapped one of its wings in a way that helped it reach random objects, rest on them and use them to lift itself up. It was, moreover, observed that the robot independently discovered (learned) movements which allowed it to hang in the air without slipping down the poles to which it was attached.[40] Thus, the robot trained itself by improving its flapping technique.[41] This experiment gave rise to the expression "if you don't know how to program a robot to fly, program it so that it can find out how to do it by itself".[42]

Another example is the escape of the Gaak robot from the Magna Science Centre in England. This instance reveals the ability of AI to learn and respond to the environment, regardless of the will of the AI developer. In 2002, the Magna Science Centre in Rotherham, England, ran a project called "Living Robots." Robots assigned to the roles of "predators" and "prey" were released into an arena of 2000 square meters. Using infrared sensors, the "prey" was looking for food indicated by light and "predators" were hunting them and draining their energy.[43] Artificial Intelligence operates by means of neural networks. The experiment had to confirm the principle of the "survival of the fittest" and reveal whether the robots (and AI as well) were able to benefit from the gained experience, i.e. to independently come up with new hunting and self-defense techniques. During the experiment, a robot Gaak that was unintentionally left unattended for 15 min managed to escape. The robot forced its way through the "prison" wall, found an exit and went to the street. It reached the parking lot where it was hit by a car.[44]

So far only primitive AI systems have been encountered in practice, but rapid technological development[45] and the current market trends, such as the Google quantum supercomputer,[46] digital copies of oneself,[47] etc., suggest that primitive AI will be replaced by AI that is distinguished by better technological properties. It is clear that Artificial Intelligence is gaining new properties and abilities that often surpass their human counterparts. This process is accelerating. Moore's law states that the capacity of computer processors doubles every eighteen months.[48] All of the foregoing is influenced by the change in human knowledge levels, or, more precisely, the continuous increase of knowledge. The level of human knowledge doubled during the period from 1750 to the twentieth century. And since 1965 the level of human knowledge doubles every five years. According to some predictions, the knowledge level shall double every 73 days, beginning in 2020.[49] As the knowledge level increases, people become incapable of controlling and generating information, thus computer programs are used as auxiliary tools. Large flows of

[37] Eric J. Sinrod, 'Perspective: Is GPS Liability Next?' (CNET NEWS2008) <http://www.mondaq.com/unitedstates/x/56280/Satellite/Perspective+Is+GPS+Liability+Next>accessed 25.09.13; Martin J. Saulen, '„The machine knows!": what legal implications arise for GPS device manufacturers when drivers following their gps device instructions cause an accident?' (2010) 44 New Eng. L. Rev.159, 160–161.

[38] For example: Promotional clip of a NAO robot demonstrates how the robot responding to other robot's question, whether it had seen a duck-shaped eraser, lied that it had not seen such an object, even though it had thrown the eraser in the trashcan a few minutes earlier. <http://www.youtube.com/watch?v=2STTNYNF4lk> accessed 05.11.13.

[39] Ibid.

[40] Peter Augustsson, Krister Wolff, Peter Nordin, 'Creation of a Learning, Flying Robot by Means of Evolution' [2002] In Proceedings of the Genetic and Evolutionary Computation Conference GECCO 02. Chalmers University of Technology.

[41] Ian Sample, 'Winged robot learns to fly' [2002] NewScientist, <http://www.newscientist.com/article/dn2673-winged-robot-learns-to-fly.html#.Us6iibS9aSq> accessed 06.11.13.

[42] Lakshmi Sandhana, 'A Theory of Evolution, for Robots' [2002] 25 November 2008 <http://www.wired.com/science/discoveries/news/2002/09/54900?currentPage=all>accessed 04.11.13.

[43] Martin Wainwright, 'Robot fails to find a place in the sun' [2002] The Guardian<http://www.theguardian.com/uk/2002/jun/20/engineering.highereducation>accessed 25.11.13.

[44] Točka (n 4) 17–18.

[45] Gordon E. Moore, 'Cramming More Components onto Integrated Circuits' (1998)86 (1) Proceedings of The IEEE 82, 82–85; Michael Kanellos, 'Moore's Law to roll on for another decade' [2003] CNET NEWS<http://news.cnet.com/2100-1001-984051.html> accessed 05.02.14; Luke Muehlhauser, Anna Salamon, 'Intelligence Explosion: Evidence and Import', In *Singularity Hypotheses: A Scientific and Philosophical Assessment*, (ed. Amnon Eden, Johnny Søraker, James H. Moor, Eric Steinhart, Springer 2012), 15–42. Web version <http://intelligence.org/files/IE-EI.pdf>accessed 25.09.13.

[46] Google and NASA Are Building the Future of AI With a Quantum Supercomputer. Available form internet: <http://gizmodo.com/google-and-nasa-are-building-the-future-of-ai-with-a-qu-507376295>accessed 25.09.13.

[47] Adee Sally, 'Digital doppelgängers: Building an army of you' [2012] <http://www.newscientist.com/article/mg21528771.200-digital-doppelgangers-building-an-army-of-you.html>accessed 25.09.13.

[48] Kanellos (n 46).

[49] Točka (n 5) 17.

information and data require more powerful computer (technological) systems.[50] Powerful computer systems lead to the growth of their advantage over humans. Therefore, the divide between human capabilities and computer systems is constantly growing.

With the rapid growth of computer capacities, their memory volume, self-training, and creative potential, the difference between the abilities of the human brain and machine intelligence is decreasing, and in this respect we may switch places in the future. Already, computer memory consists of elementary circuits, the number of which is comparable to that of nerve cell connections in the human brain (about 10 trillion), and the speed of their operation is higher than that of the human brain's (billions of operations per second). It is believed that in 10–20 years artificial and human intelligence will be equal, and subsequently Artificial Intelligence will "outgrow" human intelligence.[51]

According to Google engineer, futurist (futurologist) Ray Kurzweil, Artificial Intelligence will become superior to the human one in 2045.[52] Having defined the Law of Accelerating Returns[53] on the basis of Moore's law,[54] Kurzweil notes that faster and smarter (self-replicating and self-oganizing) chips will themselves accelerate the growth of power of computers.[55] In his book *The Singularity is Near*,[56] Kurzweil analyzes operational capabilities and the development of AI in the near future. He believes that in 2045, with the creation of superintelligence capable of self-improvement, an intelligence explosion or "technological singularity" shall occur. This superintelligence shall undoubtedly be superior to human intelligence; it will become autonomous and independent from the will of humanity.[57]

2.3. *Risks posed by AI: factors that influence the occurrence of damage*

Although superintelligence capable of self-improvement and possessing the ability to create Artificial Intelligence has not been created yet, the existing robots with integrated primitive Artificial Intelligence systems can cause damage to persons. This issue was extensively analyzed by Stephen M. Omohundro,[58] who stated that even AI with the ability to play chess may be dangerous, if not properly designed,[59] AI developed without any special precautions shall oppose its own disconnection and try to break into other technological systems in order to create a copy of itself. It is also worth noting that improperly designed AI may try to acquire resources without regard to the safety of others in order to achieve the goal for which it was designed.[60]

The examples of AI-based software presented above only confirm the potential occurrence of damage. With the ability to train itself and accumulate experience, AI can make decisions independent of the will of its developer. The escape of the Gaak robot from the Magna Science Center is an excellent example of how a robot caused a traffic accident through independent decision-making and, thereby, caused damage to another person.

With the introduction of more advanced AI systems, the likelihood of damage will only increase. AIs common in social life (i) will want to self-improve; (ii) will want to be rational; (iii) will try to preserve their utility functions; (iv) will try to prevent falsification of their operating results, i.e. counterfeit of their functional properties; (vi) will want to acquire resources and use them efficiently.[61] According to Muehlhauser and Salamon, who make reference to the theories of Stephen M. Omohundro[62]and Nick Bostrom,[63] these aspirations are only intermediate, convergent objectives that lead to the final goal for which AI has been created.[64] Thus, by achieving intermediate objectives in order to achieve the final goal, Artificial Intelligence can cause damage to third parties.

According to the analysis of the principles of the operation of AI performed by Omohundro and Bostrom, the following factors that determine the occurrence of damage (resulting in damage) can be identified:

1) The objective of AI to preserve itself in order to maximize the satisfaction of its present final goals[65];
2) The objective of AI to preserve the content of its current final goals; otherwise, if the content of its final goals is changed, it will be less likely to act in the future to maximize the satisfaction of its present final goals[66];

[50] *Digital Universe study,* iView content – Digital Universe studyiView content – Final (2010) 8–10, <http://gigaom.files.wordpress.com/2010/05/2010-digital-universe-iview_5-4-10.pdf>accessed 04.11.13; John Gantz, David Reinsel, *The Digital Universe in 2020: Big data, Bigger digital shadows, and biggest Growth in the East* (2012) 13–14; <http://idcdocserv.com/1414> accessed 08.11.13.
[51] Toločka (n 5) 19–20.
[52] Ray Kurzweil, The Singularity is near. When Humans transcend biology (NY Penguin Group 2005) 98.
[53] "The Law of Accelerating Returns" - the rate of change in a wide variety of evolutionary systems (including but not limited to the growth of technologies) tends to increase exponentially.
[54] Ray Kurzweil, *The Law of Accelerating Returns*. (2001) <http://www.kurzweilai.net/the-law-of-accelerating-returns>accessed 04.09.13.
[55] Ray Kurzweil, *The Age of Spiritual Machines* (NY Penguin Books 1999) <http://www.us.penguingroup.com/static/packages/us/kurzweil/excerpts/exmain.htm>accessed 14.09.13; Kurzweil, *The Singularity is near. When Humans transcend biology* (n 53) 85.
[56] Kurzweil, The Singularity is near. When Humans transcend biology (n 53).
[57] Ibid 179.
[58] Stephen M. Omohundro, 'The Basic AI Drives. - Self-Aware Systems.' (2008) Proceedings of the 2008 conference on Artificial General Intelligence. Amsterdam: IOS Press 483–492.
[59] Ibid 483.
[60] Ibid.
[61] Muehlhauser, Salamon (n 46) 15.
[62] Omohundro (n 48) 483–492. the programmers of the software run on such machines, the users of AI, their owners and the intelligent systems themselves.
[63] Nick Bostrom, 'The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents' In *Theory and Philosophy of AI* edited by Vincent C. Müller. Special issue (2012) 22 (2) Minds and Machines, 71–85. <http://www.nickbostrom.com/superintelligentwill.pdf> accessed 08.11.13.
[64] Muehlhauser, Salamon (n 46) 15.
[65] Ibid.
[66] Muehlhauser, Salamon (n 46) 15.

3) The objective of AI to improve its own rationality and intelligence in order to improve its decision-making, and thereby increase its capacity to achieve its final goals[67];
4) The objective of AI to acquire as many resources as possible, so that these resources can be transformed and put to work for the satisfaction of AI's final goals.[68]

The above factors that may result in damage confirm that the operation of AI is based on the achievement of goals. For example, the Gaak robot used its accumulated experience in order to survive in a changing environment. To achieve its goal, the robot escaped from the research center into a parking lot, where it was hit by a car. The Gaak robot's behavior surprised even its creator, because the AI system was not programmed to perform specific actions (Wainwright, 2002). Gaak made an independent decision to escape from the research center. Therefore, in this case, the question is: who is liable for the actions of Gaak and who must compensate for the damage caused? The actors may include: producers of the AI systems machines, the users of AI, the programmers of the software run on such machines, their owners and the intelligent systems themselves.

3. The lack of direct legal regulation in the field of liability for damage caused by AI

3.1. *General basics*

In 1996, Tom Allen and Robin Widdison argued that "soon, our autonomous computers will be programmed to roam the Internet, seeking out new trading partners – whether human or machine … At this point we must inquire whether existing contract law doctrine can cope with the new technology, and if so, how."[69] The conclusion they came to was that "neither American nor English law, as they currently stand, would confer legal status on all computer-generated agreements."[70] It means that the contract law doctrine that existed in 1996 could not cope with the damage created by technologies, thus a task for the future arose: to determine how the law should be changed.

Eighteen years passed after the article by Allen and Widdison was published; and even though contracts made by the interaction of automated message systems is recognized and legal binding,[71] a similar question arises: whether the existing legal doctrine can cope with the new, sophisticated technology and the damage made by AI, and if so, how?

An explanatory note by the UNCITRAL secretariat on the United Nations Convention on the Use of Electronic Communications in International Contracts states a general principle in article 12, defining that a person (whether a natural person or a legal entity) on whose behalf a computer was programmed should ultimately be responsible for any message generated by the machine.[72] The next Electronic Communications Convention Explanatory note section (213) of Article 12 outlines that:

> *Article 12 is an enabling provision and should not be misinterpreted as allowing for an automated message system or a computer to be made the subject of rights and obligations. Electronic communications that are generated automatically by message systems or computers without direct human intervention should be regarded as 'originating' from the legal entity on behalf of which the message system or computer is operated. Questions relevant to agency that might arise in that context are to be settled under rules outside the Convention.*

Obviously that legitimate IT transaction is concluded, but national and international law does not recognize AI as a legal person. This is precisely the reason why some authors,[73] have raised the question of whether artificial agents should be recognized as legal persons. In view of such practice, there is a question of liability for damage in the context of the legal relationship between Artificial Intelligence and its developer. Legal norms provide that damages caused by unlawful actions of another person must be compensated. Damage is one of the main conditions of civil liability, which must be proven in order to obtain redress. Both the Roman law and civil codes of various civil law tradition countries[74] provide that damages have to be compensated by the offender or a person who is responsible for the actions of the offender.

If AI would be fully autonomous (such as superintelligence), then they must be aware of their actions. If they are aware of their actions, they must be liable for their actions. AI's autonomy in law means that AI has rights and a corresponding set of duties. In law, rights and duties are attributed to legal persons, both natural (such as humans) and artificial (such as corporations). Therefore, if we seek for AI to be liable for its actions, we should ascribe legal personhood to it.[75] It means that lawmakers have to review the existing legal framework and adapt it to the changing needs of society. Currently, similar activities are being performed under the RoboLAW project.

In order to adapt the existing legal framework to the changing needs of society, it is important to take into account the different operating principles of the state and the newly

[67] Ibid.
[68] Ibid.
[69] Tom Allen, Robin Widdison, 'Can computers make contracts?' (1996) 9(1) Harvard Journal of Law & Technology 25, 28–29.
[70] Ibid 52.
[71] The article 12 of United Nations Convention on the Use of Electronic Communications in international Contracts adopted in 23 November 2005, says that A contract formed by the interaction of an automated message system and a natural person, or by the interaction of automated message systems, shall not be denied validity or enforceability on the sole ground that no natural person reviewed or intervened in each of the individual actions carried out by the automated message systems or the resulting contract.
[72] Ugo Pagallo, The laws of robots: crimes, contracts, and torts (Springer 2013) 98.
[73] Miglė Laukytė, 'Artificial and Autonomous: A Person?' (2012) *Social Computing, Social Cognition, Social Networks and Multiagent Systems Social Turn - SNAMAS 2012*, ed. G. Dodig-Crnkovic, A. Rotolo et al. AISB/IACAP World Congress 2012.
[74] For example: Article 6.263 of the Civil Code of the Republic of Lithuania, Article 1383 of the Civil Code of France, Article 823 of the Civil Code of Germany, Article 1457 of the Civil Code of Quebec.
[75] Ibid 4.

emerging network society, as discussed in Manuel Castell's *The Rise of the Network Society: The Information Age: Economy, Society, and Culture*.[76] Institutions that will supervise AI activities as well as the legal framework governing the functioning of cyberspace are based on: deliberate consistency, formality, verticality, hierarchy, static nature and stability. IT, including Artificial Intelligence systems, is based on contrary principles: it is spontaneous, constantly evolving and changing.[77] In view of the foregoing, legislative acts governing this area should be: either (i) universal in order to be effective, regardless of changes in information technology, or (ii) constantly amended in order to be effective, regardless of changes in information technology. However, the second option, i.e. constant amendment of legislative acts depending on changes in information technology, may be difficult to implement due to the static and consistent nature of operation of institutions. For this reason, priority must be given to fundamental legal norms and general principles of law. Such observations should be mentioned with respect to the RoboLAW project group, who generate the “Guidelines on Regulating Robotics,” addressed to the European Commission, in order to establish a solid legal framework for the development of robotic technologies in Europe.

3.2. The RoboLaw project

The RoboLAW project (full title: *Regulating Emerging Robotic Technologies in Europe: Robotics Facing Law and Ethics*)[78] was officially launched in March, 2012. It is funded by the European Commission for the purposes of investigating ways in which emerging technologies in the field of biorobotics (and AI as well) have a bearing on the national and European legal systems, challenging traditional legal categories and qualifications, posing risks to fundamental rights and freedoms that have to be considered, and more generally demanding a regulatory ground on which they can be developed and eventually launched.[79]

Building on the perception of a pressing need for a legal framework to accompany the development of robotic technologies, the aim of the research is to outline a comprehensive analysis of the current state-of-the-art regulation pertaining to robotics in different legal systems, in order to understand whether new regulation is needed or whether the problems posed by robotic technologies can be handled within the framework of the existing laws.[80]

The most important outcome of RoboLaw is a final report containing the “Guidelines on Regulating Robotics,” which was presented on 22 September 2014. It is addressed to the European Commission, in order to establish a solid legal framework for the development of robotic technologies in Europe. The guidelines are meant for use by the European Commission in order to respond to the ethical and practical concerns regarding the application of emerging technologies. The Guidelines on Regulating Robotics is the result of cross-border discussions, both in the sense of gathering multiple nationalities and combining multiple scientific disciplines, and of a wide dissemination activity through workshops, conferences and meetings.[81]

For all of the foregoing, the question “could Artificial Intelligence become a legal person” is still only theoretical from today's perspective.[82] While the EU-driven RoboLaw project that will promote the development of guidelines governing the operation of robotics, including AI, is being carried out, AI has no legal personality. Therefore, in litigation for damages, AI may not be recognized as an entity eligible for the compensation of damages. However, in terms of law, a situation where damages are not compensated is impossible.[83] The legal system establishes liability of those responsible for the injury, the so-called “legal cause” of the injury.[84] But if AI is not a legal entity, who is to compensate for damages caused by it?

4. The AI-as-Tool concept in AI liability terms

As mentioned above, the general principle in article 12 of the United Nations Convention on the Use of Electronic Communications in International Contracts states that a person (whether a natural person or a legal entity) on whose behalf a computer was programmed should ultimately be responsible for any message generated by the machine. This interpretation complies with the general rule that the principal of a tool is responsible for the results obtained by the use of that tool since the tool has no independent volition of its own.[85]

[76] Manuel Castells, The Rise of the Network Society: The Information Age: Economy, Society, and Culture, (Volume I, 2nd, UK Blackwell Publishers Ltd 2000), 594.

[77] Matthew Fraser, Dutta Soumitra, Throwing Sheep in the Boardroom: How Online Social Networking Will Transform Your Life, Work and World (Translator Aidas Jurašius, Eugrimas 2010), 17–21.

[78] The project funded under the 7th FP (Grant Agreement No. 289092) was launched in March, 2012, and will last for 24 months. The research is carried out by a consortium of four partners from various institutions and with different backgrounds and expertise: the Dirpolis Institute and the Biorobotics Institute of the Scuola Superiore Sant'Anna in Pisa, the Tilburg Institute for Law, Technology and Society (TILT) of Tilburg University, the School of Systems Engineering of the University of Reading, and the Department of Philosophy of Humboldt University of Berlin.

[79] Erica Palmerini, ‘The interplay between law and technology, or the RoboLaw’ In *Law and Technology. The Challenge of Regulating Technological Development* (Pisa University Press 2012), 7–8.

[80] Ibid 7.

[81] RoboLaw Regulating Emerging Robotic Technologies in Europe. *Robotics Facing Law and Ethics*. Collaborative Project FP7 GA 289092 information <http://www.robolaw.eu/RoboLaw_files/documents/RoboLaw_20121004_Brochure_final.pdf> accessed 20.11.13.

[82] Lawrence B. Solum, ‘Legal Personhood for Artificial Intelligences’ (1992) 70 North Carolina Law Review 1231, 1231.

[83] Legal norms establish the obligation to compensate damages caused by unlawful actions of other persons. Both the Roman law and civil codes of various countries provide that damages are compensated by the offender or a person who is responsible for the actions of the offender.

[84] Curtis E. A. Karnow, ‘Liability For Distributed Artificial Intelligences’ (1996) 11 (1) Berkeley Technology Law Journal <http://btlj.org/data/articles/vol11/Karnow.pdf> accessed 08.11.13.

[85] Pagallo (n 77) 98.

This provision is reflected by Ugo Pagallo, who says that strict liability in the field of contracts, rights and obligations established by AI are generally interpreted through the traditional legal viewpoint defining a robot (AI)-as-tool. It means that strict liability rules govern the behaviour of that machine, and bind the natural or legal person on whose behalf it acts, regardless of whether such conduct was planned or envisaged. In tort law, strict liability rules in the field of AI are most of the time understood by analogy with a party's responsibility for the behaviour of animals, children, employees or even ultra-hazardous activity.[86] It can be recognized as the Respondeat Superior rule based upon the concept of vicarious liability.

4.1. AI liability in the context of the respondeat superior liability theory

The Respondeat Superior (Latin: "let the master answer") rule is also called the "Master-Servant Rule."[87] This regulation was established by the praetorian law in ancient Rome. Praetor's Edict provided for cases in which a claim on obligations arising under transactions of a slave who was directly involved in commercial activities could have been made against the slaveholder. All these claims had a common title: actiones adiecticiae qualitatis, because adicitur, i.e. liability of the slaveholder secured with a claim, used to be attached to the natural obligation of the said slave. In each individual case, depending on the type of activities that created the obligation, a specific claim was made.

In the case of torts, Roman law invoked a general rule: "that upon a tort a noxal action lies against the dominus, under which he must either pay the damages ordinarily due for such a wrong, or hand over the slave to the injured person."[88] Noxal liability is entirely sui generis: its form is due to its descent from ransom and from vengeance. It has points of similarity with both direct and representative liability, and expressions are used implying one or the other according to the needs of the moment.[89]

Analysis of the legal status of AI and its liability for its actions reveals a link between the status of AI and that of slaves established in the Roman law. Both AI and the slave are not subjects of law, but rather its objects. They could not apply to courts, because only free persons could participate in litigation.[90] Assuming that the parallel between legal status of AI and that of slaves is possible, it can be stated that damages caused by the actions of AI should be compensated by its owner or AI developer—the legal person on whose behalf it acts. In Roman law, this meant that the person (head of household) responsible for persons alieni iuris (subordinate slaves), i.e. their owner, was held liable for torts committed by the slaves.

4.2. Vicarious liability for AI actions

In accordance with the provision that the AI system is a tool, the parallel with a party's responsibility for the behaviour of children or employees can be used. The following relationships are the best examples of vicarious liability: (a) liability of the principal for the act of his agent or liability of the parents for their child; (c) liability of the master for the act of his servant (example discussed above). This means that for AI's behaviour vicarious liability appears to a person on whose behalf it acts or at whose disposal and supervision AI is. It can be listed as users of AI or their owners.

The Vicarious Liability Doctrine, widely analyzed by Paula Giliker in her Vicarious Liability in Tort: *A Comparative Perspective*[91] compares common law systems and civil law systems.[92] The most common example is in a workplace context where an employer can be liable for the acts or omissions of its employees, provided it can be shown that they took place in the course of their employment. In one sense this liability is strict in that the employer cannot escape it when the mentioned conditions are met. However, the analysis commissioned in 2004 by the European Commission showed that German law constitutes an exception. It is the only one of the reviewed legal systems[93] which allows the employer to avoid liability by proving that the employee was carefully selected and controlled.[94]

However, no matter how subtle the differences between legal systems are, the basic rule of vicarious liability remains the same: responsibility which renders the defendant liable for the torts committed by another means liability is imposed on the person, not because of his own wrongful act, but due to his relationship with the tortfeasor. So the viewpoint defining robot-as-tool, in other words, means that liability for the actions of AI should rest with their owners or the users. Summarising the outcomes of the robots-as-tools vicarious liability concept, we can rely on the insights of U. Pagallo who states that:

(a) AI acting on behalf of the principal P, so as to negotiate and make a contract with the counterparty C;
(b) Rights and obligations established by AI directly bind P, since all the acts of AI are considered as acts of P;
(c) P cannot evade liability by claiming either she did not intend to conclude such a contract or A made a decisive mistake;
(d) In case of the erratic behaviour of AI, P may claim damages against the designer and producer of AI. However, according to the mechanism of the burden of proof, P will have to demonstrate that AI was defective and that such defect existed while AI was under the

[86] Ibid xiv.
[87] Ralph Dornfeld Owen, 'Tort Liability in German School Law', 20 (1) Law and Contemporary Problems 72<http://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=2623&context=lcp> accessed 15.11.13.
[88] William Warwick Buckland, The Roman Law of Slavery: *The Condition of the Slave in Private Law from Augustus to Justinian* (Cambridge, the university press 1908) 98.
[89] Ibid.
[90] Nekrošius Ipolitas, Nekrošius Vytautas, Vėlyvis Stasys, *Romanlaw* (2nd edn, Justitia 1999) 64–65.
[91] Paula Giliker, Vicarious Liability in Tort: *A Comparative Perspective* (Cambridge University Press, 2010).
[92] Ibid xlii.
[93] The analysis is limited to France, Germany, Italy, Spain, Sweden, UK and the United States.
[94] Ulrich Magnus, Hans-W. Micklitz, Final report in *Comparative analysis of national liability systems for remedying damage caused by defective consumer services* (A study commissioned by the European commission, 2004) 7, <http://ec.europa.eu/consumers/archive/cons_safe/serv_safe/liability/reportd_en.pdf> accessed 15.06.14.

manufacturer's control; and, moreover, the defect was the proximate cause of the injuries suffered by P.[95]

4.3. Strict liability arising out of the actions of AI treated as animal, product or ultra-hazardous activities

There are three main cases where strict liability applies: (a) injuries by wild animals; (b) products liability; (c) abnormally dangerous activities:

(a) There are no grounds to equate AI to an animal because the activities of AI are based on an algorithmic process similar to rational human thinking and only partially similar to instincts and senses like those of animals. It is presumed that AI can understand the consequences of its actions and distinguish itself from animals. This leads to the conclusion that we cannot apply strict liability which would be applied in cases where the damage is caused by an animal.

(b) Also in some cases it would be difficult to apply the product liability case, because AI is a self-learning system that learns from its experience and can take autonomous decisions. Thus, for the plaintiff it would be difficult to prove an AI product defect and especially that the defect existed when AI left its manufacturer's or developer's hands. It is hard to believe that it is possible to draw the line between damages resulting from the AI will, i.e. derived from self-decision, and damages resulting from product defect; unless we would equate the independent decision-making (which is a distinctive AI feature) with a defect.

In this way, there is the danger that for any AI independent decision responsibility will arise for its manufacturer and finally for the programmer—the final element of the liability chain. So in this case, the burden of responsibility would be disproportionate to that person. Too large a burden of responsibility can lead to fear on the part of the programmer not to reveal his identity in public, or otherwise it can stop the progress of technology development in official markets, moving all the programming work into unofficial markets. Assuming that such a case is not possible it is obvious that to apply product liability in the case of AI is difficult or even legally flawed.

(c) If AI was treated as a greater source of danger and the person on whose behalf it acted was declared its manager, the person could be held liable without fault. Thus the question is whether AI software systems can be recognized as a greater source of danger. There are two main theories of the greater source of danger: that of the object and that of the activities. Under the theory of object, the greater source of danger is an object of the physical world that cannot be fully controlled by a person. The theory of activities provides that the greater source of danger is certain types of activities associated with greater danger to others. Both theories imply greater danger of certain objects to persons.

Therefore, it is clear that a greater source of danger is defined as a specific object of the physical world that has specific properties. That is precisely what AI is, i.e. a specific object characterized by specific properties inherent only to it. Since AI is able to draw individual conclusions from the gathered, structured, and generalized information as well as to respond accordingly, it should be accepted that its activities are hazardous.[96] Accordingly, the AI developer should be held liable for the actions of the greater source of danger, and, in this case, liability arises without fault.

Previously discussed examples that reveal in detail the operating principles and specifics of AI confirm that AI software systems can be regarded as a greater source of danger. The example with Gaak confirms that activities of AI are risky and the risk may not always be prevented by means of safety precautions. For this reason, AI meets the requirements for being considered a greater source of danger, and the manager of a greater source of danger is required to assume liability for its actions by insuring AI.

Liability without fault is based on the theory of risk. The theory is based on the fact that a person carries out activities that he or she cannot fully control; therefore, a requirement to comply with the safety regulations would not be reasonable, because even if the person acted safely, the actual risk of damage would still remain.[97] In this case, it would be useful to employ the "deep pocket" theory which is common in the US. The "deep pocket" theory is that a person engaged in dangerous activities that are profitable and useful to society should compensate for damage caused to the society from the profit gained. Whether the producer or programmer, the person with a "deep pocket" must guarantee his hazardous activities through the requirement of a compulsory insurance of his civil liability.[98]

Another possibility is to divide responsibility among a group of persons by grafting the Common Enterprise Doctrine onto a new strict liability regime. This idea has been raised by David C. Vladeck who argues that each entity within a set of interrelated legal persons may be held liable jointly and multiply for the actions of other entities that are part of the group. Such liability theory does not require that the persons function jointly; it would be enough to work towards a common end—such as to design, program, and manufacture an AI and its various component parts. "A common enterprise theory permits the law to impose joint liability without having to lay bare and grapple with the details of assigning every aspect of wrongdoing to one party or another; it is enough that in pursuit of a common aim the parties engaged in wrongdoing".[99] This theory of liability in the context of the lack of legal regulation in the AI field requires more academic interest, discussion, and development.

[95] Pagallo (n 77) 98.

[96] *Siri* – in case of a virtual personal assistant.

[97] Dangutė Ambrasienė, Civil Law, Law of Obligations (Egidijus Baranauskas ed, Vilnius 2006) 222.

[98] *See, e.g.*, Guido Calabresi, *The Cost of Accidents: A Legal and Economic Analysis* (Yale University Press 1970) 40-41 (expanding on deep pockets and loss distribution).

[99] David C. Vladeck, 'Machines without Principals: Liability Rules and Artificial Intelligence', (2014) 89 Wash. L. Rev. 117, 149.

5. Conclusions

The emerging discipline of AI has changed attitudes towards the intellect, which was considered to be a feature exclusively inherent to biological beings, i.e. homo sapiens. The definition of AI provides that AI is any artificially created intelligence, i.e. a software system that simulates human thinking on a computer or other devices, such as: home management systems integrated into household appliances; robots; autonomous cars; unmanned aerial vehicles, etc.

AI can be defined on the basis of the factor of a thinking human being and in terms of rational behavior: (i) systems that think and act like a human being; and (ii) systems that think and act rationally. These factors demonstrate that AI is different from conventional computer algorithms. These are systems that are able to train themselves (store their personal experience). This unique feature enables AI to act differently in the same situations, depending on the actions performed before. This is very similar to the human experience. Cognitive modeling and rational thinking techniques give more flexibility and allow for creating programs that can "understand," i.e. that have traits of a reasonable person (brain activity processes).

The ability to accumulate experience and learn from it, as well as the ability to act independently and make individual decisions, creates preconditions for damage. The article identifies the following factors that determine the occurrence of damage: (i) the objective of AI to preserve itself in order to maximize the satisfaction of its present final goals; (ii) the objective of AI to preserve the content of its current final goals; otherwise, if the content of its final goals is changed, the final goal shall not be satisfied; (iii) the objective of AI to improve its own rationality and intelligence in order to improve its decision-making, allowing it to maximize the satisfaction of present final goals; (iv) the objective of AI to acquire as many resources as possible, so that these resources can be transformed and put to work for the satisfaction of the AI's final goals.

National and international law does not recognize AI as a legal person, which means that AI may not be held personally liable for damage it causes. For this reason, in the context of AI liability issues, the following principle may be applied: the general principle in article 12 of the United Nations Convention on the Use of Electronic Communications in International Contracts, which states that a person (whether a natural person or a legal entity) on whose behalf a computer was programmed should ultimately be responsible for any message generated by the machine.

In view of the foregoing, the concept of AI-as-Tool can be applied, which means that strict liability rules govern the behaviour of that machine, binding the natural or legal person on whose behalf it acted, regardless of whether such conduct was planned or envisaged. The actors may be producers of the AI systems machines, the users of AI, the programmers of the software run on such machines, and their owners.

When an AI system is understood to be a tool, one can apply vicarious or strict liability for the damage caused by AI. The vicarious liability concept comes from the respondeat superior liability theory formed in Roman law. That theory entails that responsibility renders the defendant liable for the torts committed by primitive AI; thus, liability is imposed on the person, not because of his own wrongful act, but due to his relationship with the tortfeasor AI. In the case of erratic behaviour on the part of AI, when there is damage to a third party, the person (the AI owner or user) may claim damages against the AI designer and (or) producer (product liability). However, according to AI operating principles, based on independent decision making, it would be difficult to establish a burden of proof in an appropriate manner.

Strict liability arising out of the Actions of AI can be applied for AI treated as ultra-hazardous activities. It would be ineffective to apply strict liability in cases such as (a) injuries by wild animals, and (b) product liability. In the first case there are no grounds to equate AI with an animal because AI activities are based on an algorithmic process and it can understand the consequences of its actions. This difference distinguishes AI from animals. In the second case, in a product liability case the plaintiff would find it very difficult to prove that the AI product was defective and especially that the defect existed when AI left its manufacturer or developer. AI is a self-learning system, so it can be impossible to draw the line between damages resulting from the will of AI in the process of its (self) learning and the product defect.

AI can be treated as a greater source of danger, and the person on whose behalf it acts as a manager could be held liable without fault. Thus it would be useful to employ "deep pocket" theory, which means that a person engaged in dangerous activities that are profitable and useful to the society should compensate for damage caused to society from the profit gained. A person with a "deep pocket," whether that is the producer or programmer, is required to insure against civil liability as a guarantee for their hazardous activities. Additionally, the Common Enterprise Doctrine, adapted to a new strict liability regime, can be used in this case.

Paulius Čerka, Attorney in Čerka and The Partners advocates bureau, Associate Professor; Dr., Vytautas Magnus University Faculty of Law, Department of Private Law, Kaunas, Lithuania

Jurgita Grigienė, Corresponding author: Attorney in Čerka and The Partners advocates bureau, Associate Professor; Dr., Vytautas Magnus University Faculty of Law, Department of Private Law, Kaunas, Lithuania

Gintarė Sirbikytė, Attorney assistant in Čerka and The Partners advocates bureau, Kaunas, Lithuania

REFERENCES

Allen T, Widdison R. Can computers make contracts? Harv J Law Technol 1996;9(1). 25, 25–52.

Ambrasienė D. Civil law, law of obligations. In: Baranauskas Egidijus, editor. Publishing Center of Mykolas Romeris University; 2006. p. 607.

Augustsson P, Wolff K, Nordin P. Creation of a learning, flying robot by means of evolution. In: Proceedings of the Genetic and Evolutionary Computation Conference GECCO 02. Chalmers University of Technology; 2002.

Bostrom N. The superintelligent will: motivation and instrumental rationality in advanced artificial agents. In: Müller Vincent C, editor. Theory and philosophy of AI. Minds and Machines, vol. 22 (2); 2012. p. 71–85. Special issue, http://www.nickbostrom.com/superintelligentwill.pdf [accessed 08.11.13].
Brachman RJ, Levesque HJ. Knowledge Representation and Reasoning. 2003 (Web version), http://rair.cogsci.rpi.edu/pai/library/brachmanbook7-17-03.pdf [accessed 23.05.14].
Buckland WW. The Roman law of slavery: the condition of the slave in private law from Augustus to Justinian. Cambridge: The University Press; 1908. p. 735.
Calabresi G. The cost of accidents: a legal and economic analysis. Yale University Press; 1970. p. 350.
Castells M. The rise of the network society: the information age: economy, society, and culture. 2ndvol. I. UK Blackwell Publishers Ltd; 2000. p. 594.
Čyras V. Artificial Intelligence (E-book 2008) 5 <http://www.mif.vu.lt/~cyras/AI/konspektas-dirbtinis-intelektas.pdf>[accessed 25.11.13].
David CV. Machines without principals: liability rules and artificial intelligence. Wash L Rev 2014;89. 117, 117–150.
Davis D, Shrobe H, Szolovits P. What is a knowledge representation? AI Mag 1993;14(1):17–33. http://groups.csail.mit.edu/medg/ftp/psz/k-rep.html [accessed 23.05.14].
Digital Universe study, iView content – Final (2010) <http://gigaom.files.wordpress.com/2010/05/2010-digital-universe-iview_5-4-10.pdf> [accessed 04.11.13].
Fraser M, Soumitra D. Throwing sheep in the boardroom: how online social networking will transform your life, work and world. 2010. p. 332. Translator Aidas Jurašius, Eugrimas.
Gantz J, Reinsel D. The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the East. 2012. p. 13–4. http://idcdocserv.com/1414 [accessed 08.11.13].
Gerstner EM. Liability issues with artificial intelligence software. St Clara L Rev 1993;33(1):239. http://digitalcommons.law.scu.edu/lawreview/vol33/iss1/7 [accessed 25.09.13].
Giliker P. Vicarious liability in tort: a comparative perspective. Cambridge University Press; 2010. p. 273.
Google and NASA Are Building the Future of AI With a Quantum Supercomputer. Available form internet: <http://gizmodo.com/google-and-nasa-are-building-the-future-of-ai-with-a-qu-507376295> [accessed 25.09.13].
Hallevy G. The criminal liability of artificial intelligence entities – from science fiction to legal social control. Akron Intellect Prop J 2010;4:171–201. https://www.uakron.edu/dotAsset/1139176.pdf [accessed 28.08.13].
Hayes C. Artificial intelligence: the Future's getting closer. Am LAW 1988:115.
Howe D. Artificial neural network. Dictionary.com, The Free On-line Dictionary of Computing, <http://dictionary.reference.com/browse/artificial neural network> [accessed 12.06.14].
Kanellos M. 'Moore's Law to roll on for another decade' [2003] CNET NEWS <http://news.cnet.com/2100-1001-984051.html> [accessed 05.02.14].
Karnow CEA. Liability for distributed artificial intelligences. Berkeley Technol Law J 1996;11(1). http://btlj.org/data/articles/vol11/Karnow.pdf [accessed 08.11.13].
Kurzweil R. The law of accelerating returns. 2001. http://www.kurzweilai.net/the-law-of-accelerating-returns [accessed 04.09.13].
Kurzweil R. The singularity is near. When humans transcend biology. NYPenguin Group; 2005. p. 652.
Kurzweil R. The Age of Spiritual Machines (NY Penguin Books 1999) <http://www.us.penguingroup.com/static/packages/us/kurzweil/excerpts/exmain.htm>[accessed 14.09.13] 388.
Laukytė M. Artificial and autonomous: a person? In: Dodig-Crnkovic G, Rotolo A, et al., editors. Social computing, social cognition, social networks and multiagent systems social turn – SNAMAS 2012; 2012. AISB/IACAP World Congress 2012.
Magnus, Micklitz H W, Final report in comparative analysis of national liability systems for remedying damage caused by defective consumer services (A study commissioned by the European commission 2004) <http://ec.europa.eu/consumers/archive/cons_safe/serv_safe/liability/reportd_en.pdf> [accessed 15.06.14].
McCarthy J. What is artificial intelligence?. Stanford University, Computer Science Department; 2007. http://innovation.it.uts.edu.au/projectjmc/articles/whatisai/whatisai.pdf [accessed 25.08.13].
Moore GE. Cramming more components onto integrated circuits. In: Proceedings of The IEEEvol. 86 (1); 1998. p. 82–5.
Muehlhauser L, Salamon A. Intelligence explosion: evidence and import. In: Eden Amnon, Søraker Johnny, Moor James H, Steinhart Eric, editors. Singularity hypotheses: a scientific and philosophical assessment. Springer; 2012. p. 15–42. Web version, http://intelligence.org/files/IE-EI.pdf [accessed 25.09.13].
Nekrošius I, Nekrošius V, Vėlyvis S. Roman law. 2nd ed. 1999. p. 264. Justitia.
Nilsson NJ. The quest for artificial Intelligence. A history of ideas and achievements. Web version. 2010. p. 589., http://ai.stanford.edu/~nilsson/QAI/qai.pdf [accessed 12.12.13].
Omohundro SM. The basic AI drives. – self-aware systems. In: Proceedings of the 2008 conference on Artificial General Intelligence. Amsterdam: IOS Press; 2008. p. 483–92.
Owen R D. 'Tort Liability in German School Law', 20 (1) Law and Contemporary Problems 72–79 <http://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=2623&context=lcp> [accessed 15.11.13].
Pagallo U. The laws of robots: crimes, contracts, and torts. Springer; 2013. p. 200.
Palmerini E. The interplay between law and technology, or the RoboLaw. In: Law and Technology. The Challenge of Regulating Technological Development. Pisa University Press; 2012. p. 208.
Poole P, Mackworth A. Artificial intelligence: foundations of computational agents. Cambridge University Press; 2010 (Web version), http://artint.info/html/ArtInt_47.html [accessed 23.05.14].
RoboLaw Regulating Emerging Robotic Technologies in Europe. Robotics facing law and ethics. Collaborative Project FP7 GA 289092 information <http://www.robolaw.eu/RoboLaw_files/documents/RoboLaw_20121004_Brochure_final.pdf> [accessed 20.11.13].
Rouse M. Natural language processing. 2011.. In: http://searchcontentmanagement.techtarget.com/definition/natural-language-processing-NLP [accessed 12.06.14].
Russell SJ, Norving P. Artificial intelligence: a modern approach. NJ Prentice Hall; 1995. p. 932.
Russell SJ, Norving P. Artificial intelligence: a modern approach. 3rd ed. NJ Prentice Hall; 2009. p. 1132.
Sally A. Digital doppelgängers: building an army of you. 2012. http://www.newscientist.com/article/mg21528771.200-digital-doppelgangers-building-an-army-of-you.html [accessed 25.09.13].
Sample I. Winged robot learns to fly. NewScientist; 2002. http://www.newscientist.com/article/dn2673-winged-robot-learns-to-fly.html#.Us6iibS9aSq [accessed 06.11.13].
Sandhana L. A theory of evolution, for robots. November 2008. 2002. p. 25., http://www.wired.com/science/discoveries/news/2002/09/54900?currentPage=all [accessed 04.11.13].
Saulen MJ. The machine knows!" : what legal implications arise for GPS device manufacturers when drivers following their gps device instructions cause an accident? New Eng L Rev 2010;44:159–92.
Sherman RC. The surge of artificial intelligence: time to re-examine ourselves. Implications: why create artificial

intelligence?. 1998. http://www.units.muohio.edu/psybersite/cyberspace/aisurge/implications.shtml [accessed 16.11.13].
Shiffman D. The nature of code. 2012. http://natureofcode.com/book/chapter-10-neural-networks/ [accessed 12.06.14].
Siddiqi AA. Implications of using artificial intelligence technology in modern warfare. In: Communications and Information Technology (ICCIT), 2012 International Conference; 2012. p. 30–5. http://www.taibahu.edu.sa/iccit/allICCITpapers/pdf/p30-siddiqi.pdf [accessed 25.11.13].
Sinrod E J. 'Perspective: Is GPS Liability Next?' (CNET NEWS2008) <http://www.mondaq.com/unitedstates/x/56280/Satellite/Perspective+Is+GPS+Liability+Next> [accessed 25.09.13].
Solum BL. Legal personhood for artificial intelligences. 70 N. C Law Rev 1992:1231–88.
Tadiou K M. 'Introduction to Artificial Intelligence', <http://futurehumanevolution.com/artificial-intelligence-future-human-evolution/introduction-to-artificial-intelligence> [accessed 19.05.14].
Toločka TR. Regulated mechanisms. 2008. p. 111. Technologija.
United Nations Convention on the Use of Electronic Communications in international Contracts adopted in 23 November 2005.
Wainwright M. Robot fails to find a place in the sun. The Guardian 2002. http://www.theguardian.com/uk/2002/jun/20/engineering.highereducation [accessed 25.11.13].
Welty CA. An integrated representation for software development and discovery. Dissertation. 1996., http://www.cs.vassar.edu/~weltyc/papers/phd/HTML/dissertation-1.html [accessed 23.05.14].
Wilson B. Problem solving and search in AI. 2002. http://www.cse.unsw.edu.au/~billw/cs9414/notes/mandc/mandc.html [accessed 22.03.14].