

Artificial Intelligence in Medicine and Confidentiality of Data

Tjaša Zapušek*

Abstract

This article exposes legal and moral concerns regarding the involvement of artificial intelligence in the medical profession. On one hand it puts at stake the originality of the doctor-patient relationship. Medicine has, in its core, an extremely sensitive and moral note, since it presents a unification of very personal and confident relationships, where a doctor's decisions are not only led by medical knowledge, but also by moral judgements. On the other hand, we are confronted with legal gaps concerning liability for actions performed by robots that are equipped with such algorithms that are not provided with a specific pattern when solving the task, yet they create their own, based on their observations from their environment. Even though such robots might enhance the efficiency of medical practices (through new techniques, new treatment plans, etc.), the law needs to respond and create new legal regulations that would determine the person responsible for the harm caused by such robots. In this article I have evaluated a few different options and presented my ideas for the creation of re-regulation of an employment relationship that could potentially lead us to the desirable aim, to get the legislation that would protect individuals in the case of harm when the latter is committed by robots.

Keywords: Artificial Intelligence, Medicine, Algorithms, Ethics, Robot Doctors, Employment Relationship, Personhood

* Tjaša Zapušek is enrolled into the Postgraduate Bologna Program at the Faculty of Law at University of Copenhagen, Denmark.

I. Introduction

The protection of personal data and respect for private and family life are two very fragile yet extremely important fundamental rights. Due to its fragile nature, the European Parliament insists on the need to strike a balance between maintaining and enhancing security on one hand and safeguarding data protection and privacy of individuals on the other. An awareness of the fragility of these particular areas requires a systematic legal approach in order to continue data protection in the digital age. This article focuses on the recently emerging privacy issues regarding robot doctors. The presence of artificial intelligence and robots in the workplace is not new. Manufacturers have been using automation in their factories for many years, but recently, it has moved further into mainstream consciousness. Moreover, the involvement of artificial intelligence in the medical profession (working process) will call into question the efficiency of existing labour law regulations. The percent of the digital workforce is far from marginal and we can undoubtedly expect its increase in a not-so- distant future. Inner reorganization of the latter is crucial, since modern machines will no longer be considered simple tools in the hands of their user. Therefore, who will be held liable for the harm caused by robots' unforeseeable actions? In my recommendations I have exposed the legal barriers for the imposition of legal personhood on robots and provided the framework for legal reorganization of labour law that could, through the robots' inclusion in *employer' substrate*, provide the answer on what seems to be, at first glance, an unanswerable question.

II. Moral Foundations of Medicine

The Hippocratic Oath, and especially its moral inheritance, is widely invoked in modern medical culture and presents a so-called ethical path for the modern medical profession.¹ It got the name after the Greek physician Hippocrates, who is known as

¹ Fabrice Jotterand, "The Hippocratic Oath and Contemporary Medicine: Dialectic between Past

the father of medicine.² The Oath begins with an expression of devotion to the Gods and Goddesses, which leads us to the conclusion that the ability to heal was recognized as something that is so powerful that it verges on the supernatural:

I swear by Apollo Physician and Asclepius and Hygieia and Panacea and all the gods and goddesses, making them my witnesses, that I will fulfill according to my ability and judgment this oath and this covenant (...)

(...) What I may see or hear in the course of the treatment or even outside of the treatment in regard to the life of men, which on no account one must spread abroad, I will keep to myself, holding such things shameful to be spoken about.³

This philosophy did not wholly infiltrate into the modern perception of medicine. With the increase of artificial intelligence and its inclusion into the medical profession, legislators will sooner or later be confronted with the very important question of how to design laws in order to preserve the moral foundations of medicine. Inclusion of advanced technology is virtually inevitable, since it will come as a result of more and more important socio-economic factors, such as dependence on the instructions provided by the government. For instance, doctors in Slovenia in Pomurje region are confronted with limitations regarding the time spent per patient (7 minutes)⁴. This is an important aspect that already revolutionizes the application of core moral principles of medicine as a profession. “Cost containment appeared suddenly as a ‘moral obligation’ imposed on the physician. This means that the physicians are no longer exclusively committed to their patients but also dependent on

Ideals and Present Reality,” *Journal of Medicine and Philosophy* 30 (2005): 110.

² Erwin Ackerknecht, *A Short History of Medicine* (Baltimore: The Johns Hopkins University Press, 1982), 55.

³ Ludwig Edelstein, *The Hippocratic Oath: Text, Translation, and Interpretation* (Johns Hopkins Press, 1943).

⁴ POMUREC, http://www.pomurec.com/vsebinsa/41331/Posvet_o_zdravju__nas_zdravnik_ima_za_obravnavo_pacienta_na_voljo_le_sedem_minut__v_Evropi_pa_pol_ure.

and controlled by the social institutions.”⁵ To conclude, governments want to subordinate the medical profession to economic efficiency, meaning that through the inclusion of advanced technologies could achieve a financially optimal outcome (reduction of sick leave, longer working time etc.).

Today, the relevance of the Oath to modern medicine is more or less symbolic since it contains some obsolete duties, such as commitment to the teacher and his family, sharing money with him when he needs it, prohibition of abortion, etc. However, some of the principles were included into modern national laws (prohibition of euthanasia, duty of confidentiality) and do present an obligation for medical practitioners. As the Oath says, “If I fulfill this oath and do not violate it, may it be granted to me to enjoy life and art, being honored with fame among all men for all time to come; if I transgress it and swear falsely, may the opposite of all this be my lot.”⁶

Doctors and other medical professionals, practicing the ideas of Western medicine, recognize it as a “special ethical contract for their conduct,” which is often expressed in the saying “*Primum non nocere*.”⁷ However, even though we are allowed to claim without any restraint that original principles are, according to the currently accepted medical practices, a little bit out of date, we should not overlook the Oath’s moral evaluation of medicine as a profession. The latter has undergone some very radical shifts, yet the moral note remains in the spirit of Ancient Greece.

The world was, is, and will always be a living place of various types of traditions, cultures and notions of what can be considered as morally acceptable in a particular society. These differences are mirrored in the activities and competing practices (such as different legal traditions and different medical practices), not only of an individual but also of a society he belongs to. “Oaths do not compel ethical behavior, but they are human instruments that are crafted to sensitize the reader to moral moments and choices.”⁸ At this point, we should ask ourselves, are the decisions, even the risky

⁵ Jotterand, 116.

⁶ Edelstein.

⁷ Steven H. Miles, *The Hippocratic Oath and the Ethics of Medicine* (Oxford University Press, 2004), 50, 143.

⁸ Miles, 172.

ones with (technically) an optimal outcome, always the best decisions? The answer would probably be affirmative, if it was answered by economists, businessmen, etc. Here they directly put on the line potential financial gain, yet however, indirectly they put at stake the quality of lives of many employees, their families, their possibilities to offer their children quality education, and sufficient healthcare treatment, to name a few examples.

When the potential outcome, derived by our risky decisions, is not directly connected with finances, but with the highest human values such as health, our perspective will change. Medicine has, in its core, an extremely sensitive and moral note, since it presents a unification of very personal and confident relationships. With every single treatment, patients' lives are consciously put in the hands of doctors, people they do not know. The only thing they can rely on is first and foremost the belief that the doctor's decisions will not only be led by his medical knowledge, but also by his moral judgements.

III. Medical Confidentiality And Data Protection

Until the sixties, most of the Western countries did not recognize the existence of health law and bioethics as independent disciplines.⁹ The most notorious feature of the moralistic-paternalistic model that prevailed in the previous century was that:

...the medical profession was amoral, but normative was implicit in medical practice rather than being extensively elaborated by lawyers and ethicists. Medical ethics was the ethics of good medical practice, of being a good doctor. For this there was no elaborate body of guidelines and rules, neither in moral philosophy, nor in law. Theoretical or philosophical reflection on medical issues usually did not address the public at large.¹⁰

⁹ Michael Freeman, *Ethics and Medical Decision-making* (Ashgate Dartmouth, 2001), 48.

¹⁰ *Ibid.*, 49.

In the majority of cases, patients were not provided with all the details regarding their diagnosis, especially if the prognosis was grim. Doctors used to enjoy a high level of discretionary power, especially in the field of psychiatry, which was overwhelmed with the paternalistic way of thinking. The “patient’s best interest” presented an ideal that served as a guideline for a doctor’s final evaluation about non-voluntary treatments and institutionalization.¹¹ Nowadays, the law has, in comparison with the paternalistic model, elaborated a body of rules that serves as a safeguard against non-voluntary treatment. Thus, the latter is not (exclusively) subject of a doctor’s evaluation for what counts as “patient’s best interest,” yet there exist written requirements and steps that needs to be fulfilled before a doctor is allowed to perform non-voluntary treatment.

Informational privacy relates to various kinds of information regarding an individual person. In addition, an information about an individual’s health condition belongs, due to its sensitive nature, to the private sphere and represents the core of informational privacy.¹² In Europe the right to privacy, including the protection of sensitive data related to health, is protected both in the Data Protection Directive (95/46/EC)¹³ and the European Charter of Fundamental Rights and Freedoms (Articles 6 and 8).¹⁴

The Data Protection Directive protects different categories of personal data, where those recognized as particularly sensitive enjoy additional attention and are exposed to stricter protective measures.¹⁵ According to the second paragraph of the Article 8 of the Data Protection Directive, personal data must be used fairly for special purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Additionally, everyone has the right to access the data concerning him or her, and the right to have it rectified. The abovementioned consent does not need to be obtained in situations when the processing of data is necessary for “the purposes of preventive medicine, medical diagnosis, the provision of care or

¹¹ Ibid.

¹² Brigit Toebe, Mette Hartlev, Aart Hendriks and Jane Rothmar Herrmann, eds., *Health and Human Rights in Europe* (Intersentia, 2012), 77.

¹³ European Parliament, *Directive 95/46/EC of the Council of 24 October 1995* (on the protection of individuals with regard to the processing of personal data and on the free movement of such data).

¹⁴ Charter of the Fundamental Rights of the European Union art. 6,8, 2010 O.J. C 83/02.

¹⁵ Toebe et al., 277.

treatment, or the management of healthcare services, and where the data processed by a health professional who is bound by professional obligation of secrecy, or by an equivalent duty of confidentiality.”¹⁶

Patients’ use of healthcare services has as a “necessary evil” the exposition and revelation of their private lives, since in order to be correctly diagnosed, they will have to not only expose their bodies and thoughts but also reveal their healthcare status, history of past treatments, family health history, etc. One of the cornerstones of medical ethics that have been developed through the centuries is the imposition of a duty of confidentiality on doctors and other healthcare professionals with the main purpose of protecting a patient’s right to privacy. From the philosophical aspect, we would draw the wrong conclusion if we claimed that the value is in the secret itself, because in fact a value supposes keeping the secret. *Ratio legis* of such “keeping” is in its essence a guarantee to maintain the privacy of personal information that has been confided to doctors, priests, lawyers, psychiatrists, etc. It is important to note that patients’ right to privacy can also be at issue with regard to the activities such as storage, collection, alteration, disclosure and erasure of healthcare data.¹⁷ Thus, due to their importance, each one of them needs a justification. According to Toebes et al. in their report regarding the abovementioned activities, “In general, data protection law puts an emphasis on the right to informational self-determination, and as a main rule it is necessary to obtain express consent to the processing of personal data.”¹⁸ However, consent need not be obtained when the conditions, prescribed by the Article 8 of European Convention on Human Rights (ECHR),¹⁹ are met. Public authorities cannot interfere in the enjoyment of the right to respect private and family life, except when such interference is in accordance with the law and necessary in a democratic society; in the interests of national security, public safety or the economic well-being of the country; for the prevention of disorder or crime; for the protection of health or morals; or for the protection of the rights and freedoms of others.²⁰

¹⁶ Ibid.

¹⁷ Ibid., 139-140.

¹⁸ Ibid., 140.

¹⁹ Council of Europe, *The European Convention on Human Right* (Strasbourg: Directorate of Information, 1952), Art. 8.

²⁰ Council of Europe, art. 8 (2).

In other words, the law allows individuals, pledged to secrecy, to bypass the provisions of Article 8(2) of the ECHR. The abovementioned paragraph states that any disclosure must be in accordance with the law and have legitimate purpose; moreover it has to be proportionate in accordance with the law and necessary pursuant to democratic society. Therefore, in order to justify a disclosure, some specific criteria need to be met. A patient must always be informed about the disclosure of information and consent must be obtained before its revelation or, in a situation when confidential healthcare information is used or disclosed without prior consent, it “should clearly serve one of the purposes specified in international human rights law as being a legitimate limitation on the right to privacy; and “such disclosures must also meet the criteria of being proportionate to the legitimate aim of the disclosure”; and must be “in accordance with (domestic) law.”²¹

The collection of patient’s data, laid down in medical files, is a prerequisite for the assurance of quality, sufficiency and continuity of proper medical treatment. Doctors and other healthcare professionals need to have access to the history of an individual’s medical treatments, since it enables those involved in a patient’s treatment to exchange information, opinions and knowledge. It is presumed that this kind of communication is in accordance with the patient’s will and interest, and does not violate any right to privacy and confidentiality.²² Therefore, it might be claimed that a permission for recommended treatment, provided by an individual, also implies consent to the communication between doctors and other health professionals. All the sensitive information that circulates between them must stay within the “circle of confidentiality” and moreover, the reckless revelation of data is prevented by an imposition of the obligation of professional secrecy.²³ However, as it has been emphasized by the ECHR in *I v. Finland*, the bare fact that a person is qualified as a doctor or other health professional does not give him or her the power to access every medical file he or she wants.²⁴ In the abovementioned case, the court concluded

²¹ Roy McClelland, “European Standards on Confidentiality and Privacy in Healthcare,” www.eurosocap.org.

²² Toebe et al., 140.

²³ Ibid., 141.

²⁴ ECtHR 17 July 2008, *I v. Finland*, no. 20511/03.

that medical files were “not sufficiently protected” against unauthorized disclosure to healthcare professionals who were not involved in the medical treatment.²⁵ The latter statement raises second thoughts when we imagine that the medical treatments could possibly be performed by the new generation of robots.

IV. Artificial Intelligence in Medicine

The protection of personal data and respect for private and family life are two very fragile yet extremely important fundamental rights. Due to their fragile natures, the European Parliament insists on the need to strike a balance between maintaining and enhancing security on one hand and safeguarding data protection and individual privacy on the other. Exploring the limitless capabilities of digitalization seems very tempting from various aspects. The digital era could be described as a synonym for a new, shiny wonderland. Edward Snowden’s revelations of mass surveillance have emphasized the potential threat to the privacy of communications and personal data, which calls into question the ability to protect anonymous sources. The awareness of the fragility of this particular area requires a systematic legal approach in order to continue its protection in the digital age. This article focuses on recently-emerged privacy issues regarding humanoid robots, and robot doctors in particular. The presence of artificial intelligence and robots in the workplace is not new. Manufacturers have been using automation in their factories for many years, but recently, it has moved further into mainstream consciousness.

In the light of technological developments, the issues regarding data protection in the field of healthcare reach new dimensions and pose new, controversial questions waiting to be answered. The increasing autonomy of robots strives for a comprehensive and systematic legal approach, since sooner or later robots will no longer be considered simple tools in the hands of the owner, the user, or the manufacturer. Current national legal frameworks do not presume the robot’s liability for acts or omissions that cause damage to third parties. Therefore, in this kind of

²⁵ ECtHR, paragraphs 38-49.

situation, the producer was liable for a robot's acts or omissions, if he could have foreseen and avoided the robot's harmful behavior. Such anticipation, as we will see below, will literally be impossible, since the new generation of robots will be equipped with algorithms that enable the machine to act on the basis of its observations of the surrounding area. According to the Council Directive 85/374/EEC,²⁶ dated 25 July 1985, non-contractual liability can only cover damage caused by the robot's manufacturing defects and on condition required by Article 4 of the mentioned Directive,²⁷ in that the injured person is able to prove the damage, the defect in the product, and the causal relationship between defect and damage. This legal framework will not be sufficient to cover the damage caused by robots of the new generation, since their technological improvement will equip them with the ability to learn from their own experiences, to perform activities and tasks that used to be exclusively attached to human beings, and to interact with the outside world in a unique way.²⁸ In the European Parliament's recommendation²⁹ about the creation of specific legal status for robots, it emphasized scientific achievements in the area of artificial intelligence, especially the uniqueness of medical robots. Pursuant to the recommendation medical robots would have to go through sufficient training for doctors and care assistants in order to provide the highest degree of professionalism.

For years, robots' involvement in medical procedure has been limited to "robotic assisted surgeries."³⁰ However, current scientific achievements present a completely different scenario, where medical surgeries can be performed one hundred percent autonomously by a robot. Smart tissue autonomous robots, also known as STARs, are currently able to perform only the simplest medical procedures, yet it might change in the near future. It has unique and sophisticated equipment that can independently target soft tissue. According to *Science Translational Medicine*, the STAR can also

²⁶ European Parliament, *Council Directive 85/374/EEC of 25 July 1985* (on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products).

²⁷ *Ibid.*, Article 4

²⁸ European Parliament, Draft Report with Recommendation to the Commission on Civil Law Rules on Robotics, 31.5.2016, 2015/2103 [INL], 6.

²⁹ European Parliament.

³⁰ Cohen Hiyaguha, "Robots Replacing Surgeons?" *Neutral Health Blog*, May 28, 2016, <https://jonbarron.org/doctors-and-drugs/robots-replacing-surgeons>.

create its own treatment plan. For instance, the results of the first experimental procedure on animal tissue indicate that robots performed better than their human counterparts. The subjects of an evaluation were the number of mistakes, needle placement, stitch spacing and tension and also the potential for the finished seam to leak. However, a human outshines a robot in the time needed for a completed medical procedure.³¹ The researchers responded that STAR robot had been deliberately programmed to work slowly in case something had gone wrong during the surgery.³² It has been declared that the purpose of STAR robot is to expand Western medicine's capabilities and capacities through cleverness, enhanced vision and complementary machine intelligence in order to improve surgical procedures and achieve better outcomes.³³

Advanced robots were not only successful at performing physical operations, yet they were also included in much more sensitive therapeutic procedures concerning mental disabilities. On March 7, 2016, scientists in Singapore introduced "Nadine," an emotionally intelligent humanoid robot who can feel, think, and recognize people. Researchers at Singapore's Nanyang Technological University provide her with the ability to recall previous conversations with individual persons. According to Nadia Thelman, the Director of the institute for Media Innovation, "As countries worldwide face challenges of an aging population, social robots can be one solution to address the shrinking workforce, become personal companions for children and the elderly at home, and even serve as a platform for healthcare services in future."³⁴ Thelman predicts that Nadine's ability to interact with people in such unique way will help individuals suffering from dementia, mental diseases or autism.³⁵ The latter diagnosis is characterized by atypical behavior as well as difficulties regarding social communication.

³¹ It has to be taken in the observation that time is of the essence in surgical procedures, since longer exposure to anesthesia could expose a body to stressful conditions that could potentially lead to harmful consequences.

³² Hiyaguha.

³³ Charlie Sorrell, "An Autonomous Robot Surgeon Just Successfully Sewed Up a Pig." May 9, 2016, <http://www.fastcoexist.com/3059594/an-autonomous-robot-surgeon-just-successfully-sewed-up-a-pig>.

³⁴ Dipo Faloyin. "Scientists Have Built a 'Social' Robot with a Personality," *Newsweek*, December 20, 2015, <http://www.newsweek.com/scientists-develop-social-robot-has-personality-409993>.

³⁵ "Emotionally intelligent robot comes to life 7 March 2016," YouTube video, 2:17, posted by: "tNEWS," July 3, 2016, https://www.youtube.com/watch?v=pXg33S3U_Oc&t=4s.

The Center for Disease Control and Prevention (CDC) has revealed that according to observation, “1 in 88 children and an estimated 1 out of 54 boys in the United States have ASD. That is 78% increase since the latest CDC report in 2009.” The percentage increase should not be overlooked, since at this moment a number of families are not financially able to provide their children with optimal medical treatment. The latter typically involves many hours of child-qualified therapist interaction, whose services are not available in a majority of communities.

Because the prices of the above-mentioned treatments in the majority of cases exceed the financial resources of families, high-quality therapies are inaccessible to the wide ASD population. As a response, scientists have developed a robot-mediated therapeutic approach that enables a repetitive practice of social orientation skills like joint attention skills (JA). Roughly speaking, it presents “a triadic exchange in which a child coordinates attention between a social partner and an aspect of the environment. Such exchanges enable young children to socially coordinate their attention with people to more effectively learn from others and their environment.”³⁶ It has been observed that people diagnosed with ASD preferred robot-like characteristics over that of humans and non-robotic toys. Scholars also acknowledged that in some cases individuals respond faster to robotic movements than human ones. Moreover, scientists noticed that individuals diagnosed with autism look at the robot therapist 27.65% longer than the human therapist.³⁷ Results have shown significant preferential behavior when interacting with robots, in comparison with human therapists. The success of the innovative medical alternative has not been just claimed, yet it was proven. According to the percentage of successful trials, scholars confirmed that the robot was able to perform the task with success rates similar to that of the human therapist.

³⁶ Esubalew T. Bekele, Uttama Lahiri, Amy R. Swanson, Julie A. Crittendon, Zachary E. Warren, and Nilanjan Sarkat, “A Step towards Developing Adaptive Robot-mediated Intervention Architecture (ARIA) for Children with Autism,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21, no. 2 (March 2013): 289.

³⁷ Bekele et al., 295.

V. Types of Algorithms

In the course of a robot doctor's working process, the machine will be confronted with different assignments that could be solved through a well-defined sequence of steps, which is described in the mathematical world through the synonym *algorithm*. In the text below, the differences between four decision-making processes are presented as a consequence of four different types of algorithms, in order to visualize why we need different legal approaches for different kinds of assignment.

DETERMINISTIC ALGORITHM³⁸

Patients' medical files serve as the robots' workbook, enabling them to put their technical knowledge, stored on their internal discs, into practice. In the case of a deterministic algorithm, we will always find someone liable for a robot's unjustified revelation of medical files and records of treatments, because files (stored on robots' discs) were clearly insufficiently protected. This algorithm is constructed to prevent the machine from making autonomous decisions, creating new patterns and consequently creating unpredictable outcomes:

The most widely used clinical robotic surgical system includes a camera arm and mechanical arms with surgical instruments attached to them. The surgeon (a human that control the operation and tells the robot what to do) controls the arms while seated at a computer console near the operating table. The console gives the surgeon a high-definition, magnified, 3-D view of the surgical site. The surgeon leads other team members who assist during the operation.³⁹

As emphasized, robots equipped with deterministic algorithms are not capable of making their own decisions in evaluating whether particular information obtained by a

³⁸ Mireille Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar, 2016), 23.

³⁹ Mayo Clinic, "Robotic Surgery," *Mayo Clinic*, [surgery/basics/definition/prc-20013988](http://www.mayoclinic.org/robotic-surgery/basics/definition/prc-20013988).

patient should be revealed (for public security, public interest, etc.), since they follow predefined paths. The major significance that distinguishes the first group of machines from the others is that the final action/machine's response/steps have already been predicted. Without a single doubt it is easier to draw parallels between such robots and simple tools than with actual doctors. Here, robots more or less serve, from a data protection perspective, as a digital recording station. A revelation of the data comes as a consequence of the machine's malfunction, negligent storage, or illegal action by an incompetent or malicious person who does not have the access to these files.

*SUPERVISED LEARNING ALGORITHMS*⁴⁰

More advanced forms of algorithms can enable machine learning. In the case of supervised learning algorithms, a machine is fed a so-called training set by human supervisors, which consists of particular data and predefined patterns that provide some sort of a background on what counts as a desired and satisfactory solution/outcome. Regular training may equip the machine with the ability to recognize new data, while using enhanced pattern recognition skills. As a consequence, the agent may come up with similar or even better solutions than those that would have been provided by its human supervisor, if he had concluded the task by himself.

*UNSUPERVISED LEARNING ALGORITHMS*⁴¹

The other advanced form of algorithms is an unsupervised learning algorithm, which in comparison with a supervised algorithm, makes correlations between obtained data without previously hypothesizing them. Without the provision of a training set, consisting of data and predefined patterns, the machine cannot familiarize itself with information and what counts as a desirable solution. Such machines are capable of mining new data and creating novel, unexpected patterns.

⁴⁰ Hildebrandt, 24, 25.

⁴¹ Ibid.

MACHINE BASED ON MULTI-AGENT SYSTEM ⁴²

The fourth type of machine decision-making processes is significant for the higher degree of unpredictability of a machine's actions. The multi-agent system is a result of interaction between deterministic algorithms and agents based on machine learning. These interactions do not rely on predefined behaviors, directed from a "central point of authority."⁴³ The outcome is that an agent's emergent reactions, which contribute to solving a problem, are unforeseeable for the program. The complexity of interactions between unsupervised recognition of the environment by individual agents will reach the level where human programmers will not manage to predict the appearance of undesirable consequences or to repair potential bugs or dangerous incompatibilities between the systems. The system of the interactions within a machine's environment can be described by the term "global agency."⁴⁴ In multi-agent systems, an individual agent is not capable of recognizing and analyzing the whole environment. However, the analysis of the latter is a basis for the final, optimal reaction. For this reason an individual agent, working within the multi-agent system, can rely first and foremost on its individual yet limited knowledge. The interactions of individual analyses and recognition of the data obtained from the environment may result in unique self-organization forms that are unforeseeable or could be even undesirable from our perspective. The aspect of non-foreseeability plays a key role when dealing with legal question regarding liability.

Confrontation with the last three groups of algorithms (supervised learning algorithm, unsupervised learning algorithm and multi-agent system) that are significant due to the possibility of unpredictable decisions, leads us to the conclusion that the unjustified revelation of data may come as consequence either of insufficient protection or a robot's "intentional" and unpredictable revelation. In the case of human doctors, the disclosure of clinical records cannot be treated as illegal if there exists public interest that overwhelms the duty to keep the information secret. It is up to the doctor to evaluate all relevant circumstances and to decide whether he or she

⁴² Hildebrandt, 26, 27.

⁴³ Ibid.

⁴⁴ Hildebrandt, 26.

will reveal an important piece of information or clinical record. When we try to compare human doctors and robot doctors, we recognize that the main issue is a difference in the ability to make a conscious evaluation whether this particular information threatens public values and consequently public interests. The importance of doctor's careful evaluation mirrors national legal regulations, where unjustified betrayal may lead to prosecution by national authorities. Unjustified betrayal of professional secrecy presents serious interference in basic human rights and consequently invokes criminal liability. The act of betrayal presents any intentional act (not a negligent act) of an individual (active or passive) that enables a familiarization of an ineligible person with a patient's private data. This clearly cannot be the case with robot doctors, due to the lack of intent and what is even more important, due to the lack of consciousness.

VI. Legal Personhood

The question I am trying to answer in this paper is how to regulate advanced robots and how to legally evaluate their actions and potential harms derived by their actions. As it has been described in previous sections, existing national laws regulate only the machines that serve as a simple tool and not in their advanced forms. Now, the European Parliament has an intent to create a legal basis through the attachment of legal personhood that will allow robot-doctors to integrate into a working process, where they will perform the same tasks as humans and consequently bring into question the effectiveness of basic human rights.

First, I want to critically evaluate the potential option of imposition of legal personhood on robots. It does not matter how they are created and structured—legal systems and laws as such are incapable of self-realization. Their *spiritus agens*—motive force—is a legal person, or the holder of not only legal actions, but also legal changes and legal consequences.⁴⁵ In the world of law, an individual can be described as a “legal actor” to whom a “legal system” acknowledges a “legal role.”

⁴⁵ Marijan Pavčnik, *Teorija prava* (GV Založba, 2011), 136.

The “legal role” we play is composed of the rules and obligations that enhance us with the ability to enter legal relationships with the main purpose of achieving acceptable legal aims. At the end of the day, a legal actor is always a human being who plays the acknowledged legal role either as a physical person or as a member of legal entity, a non-human entity that is treated as a person for certain legal purposes. Human beings as the final decision maker must be aware of the consequences of his or her actions. Decisions with all of the attached legal consequences return to him or her afterward, with all of the attached rights and obligations —*Hominum causa omne ius constitutum*.⁴⁶ The main reason for the assignment of legal personhood is a value-based decision to assign certain rights and obligations to the legal actor.⁴⁷ The assignment of such quality acknowledges legal actors as holders of rights and obligations, capable of making decisions and at the same time being aware of legal consequences. From this point of view, the assignment of legal personhood can always be described as “non-natural.”⁴⁸ However, in order to become a legal actor, there needs to be not only a human, but also his or her personal decision behind “the mask.” Meaning, the legislatures would never be able to assign legal personhood to legal entity which in this case serves as “the mask,” if there were not real people with ambitions and visions behind it. The same can be claimed for robots. The attachment of legal personhood on robots would undermine the idea of legal personhood. Even though we will get “the mask”/robot, there will not be a person standing behind robot’s actions.

VII. Liability of the Producer

Someone might claim that in this case the burden of liability should be attached to the producer instead of experimenting with the attachment of legal personhood. For example, Council Directive 85/374/EEC (Directive)⁴⁹ sets the rules concerning

⁴⁶ Ibid., 137.

⁴⁷ Ibid.

⁴⁸ Ibid., 138.

⁴⁹ European Parliament, *Council Directive 85/374/EEC of 25 July 1985* (on the approximation of

producer's liability for a defect product. According to the Directive, term "product" presents a material thing that is run by electricity. The product can be described as defective when it does not provide the safety that a person is entitled to expect when he/she takes into account all of the relevant circumstances, such as presentation of the product, its use and the time when the product was put into circulation.⁵⁰ Article 4 of the mentioned Directive requires from an injured person the proof about the damage, the product's defect and the causal relationship between the two.⁵¹ Pursuant to Article 3 of the Directive, the word producer means the manufacturer of a finished product, of a raw material, of a component part and also a person who attaches his name, trademark or some other distinguishing feature on the product, as a "label," presenting him as its producer.⁵² In a situation when a producer is unidentifiable, any supplier of the product shall be treated as product's producer. However, neither the producer nor the supplier (in a case of un-identifiability of a product) can be held liable for the damage/injury caused by the defect product if he proves (among the exculpatory reasons, listed in the Article 7 of the Directive) that the product's defect that lead to the damage did not exist at the time when it was put into circulation. When applied to the case of robot doctors, since we are talking about machines equipped with such algorithms that can, through observation of the surrounding areas, learn from their own experiences and create their own patterns about how they will conclude a particular task. Here, the producer can reasonably claim that the defect which caused the harm did not exist at the time when the machine was put into circulation.

VIII. Re-regulation of Employment Relations

In the two previous sections I have evaluated two potential solutions: the first was recommended by European Parliament, where I concluded that an imposition of legal

the laws, regulations and administrative provisions of the Member States concerning liability for defective products).

⁵⁰ *Ibid.*, Article 6.

⁵¹ *Ibid.*, Article 4.

⁵² *Ibid.*, Article 3.

personhood on robots would shake the cornerstones on which this legal fiction was built. The analysis of existing rules regarding product liability, I have observed that they are clearly not sufficient. In this section I will present the other potential solution, my creation of re-organized employment relationships. With the latter, the existing law will be able to find a person responsible for robot's unpredictable actions and therefore effectively protect the harmed party.

Employers are the ones who organize and coordinate working processes that fall within the scope of their main activity. The organization of working processes requires, besides general coordination of his employees, also the provision of sufficient working space, machines, necessary facilities and other conditions crucial to maintain a good and healthy working environment. The nature of his work gives an employer a great sense of autonomy, since his organization and coordination affects how his employee carries out every particular action he is assigned. When creating a way in which a particular action will be carried out, he bears in mind the risks and possible consequences of different options. The inclusion of advanced types of robots into the working process presents to the employer a sense of uncertainty, since he cannot predict the risks and possible consequences. On the other hand, at the moment the inclusion of advanced types of robots does not impose legal risks, since there is no legal framework that would predict who the responsible person for the harm caused by robots would be.

When we ask ourselves, who is responsible for a harm occurred during or in relation to an employer's working processes, we should keep in mind that in relation to a third party, an employer will be held accountable for all harmful commitments by his *personal substrate*. In a situation when harm comes as a consequence of the operation of a machine, the law recognizes an employer as a strictly liable person, since it presumes that the machine was used by an employee, hence by someone who is a part of an employer's *personal substrate*.⁵³ To conclude, the law does not regulate emergent (not automated) actions, performed by the machine, yet it treats the machine as something that needs to be used (as a simple tool). Consequently, if we are using a machine, we stand behind its action. What is important here is the aspect

⁵³ N. Plavšak, M. Juhart, and R. Vrenčur, *Obligacijsko pravo* (GV Založba 2009), 568.

of predictability—hence if we give the machine a specific command, we know exactly what to expect.

Strict liability, also known as absolute legal responsibility, is the legal responsibility for damages and injuries that does not require guilt as a prerequisite for its imposition. An employer can be relieved of his strict liability for a harm that occurs as a consequence of the operation of the machine (in existing legislation recognized as simple tool), if he proves the actual reason for a harm is an action or an occurrence that fall outside his sphere.⁵⁴

In order to define an individual action as an action that falls outside the employer's sphere, it must cumulatively satisfy two requirements. First, the action must be uncontrollable, meaning that the employer was not able to prevent its occurrence nor deter its negative consequences. Second, the action must be unpredictable, in other words an employer was not able to anticipate its occurrence or, more accurately, an employer was not obliged to expect it.⁵⁵

Therefore, according to current legislation, the employer will always be relieved of his strict liability for harms that occur as a consequence of the operation of the machine, since he will neither be able to anticipate the occurrence of harm (even the best computer scientists do not have an insight into emerging correlations in advanced algorithms) nor will he be obliged to expect it.

I suggest that we split the workforce in two different groups. The first would be the digital workforce and the second will be human workforce. The human workforce is capable of autonomous, yet still unpredictable decision-making, just as it seems to be the case (in the sense of unpredictability) for the machines of new generation. This characteristic prevents us from using existing legal rules. The attachment of the word “mechanical” next to the word “workforce” enables *the adjustment of legal rules* (or the permission of discriminatory rules, for instance that robots will not need sick leave, etc.) regarding the human workforce and to use them for the machines of new generation that are equipped with the ability to perform activities and tasks that used to be exclusively attached to the human workforce. To sum up, the aim is to

⁵⁴ Ibid., 580.

⁵⁵ Ibid., 581.

include the mechanical workforce into employer's *substrate* in order to protect a third party against the situation when the law does not predict an individual responsible for harmful consequences.

Now, I will present how this legal regulation would work in practice. When an employee/human workforce causes harm to the third party, we should first look at the strict liability rules. Here, certain prerequisites need to be met in order to treat a harmful behavior committed by an employee as a behavior of an employer. The first requirement is that an employee needs to be a part of employer's *personal substrate* and the second one requires that an employee, as a part of the *personal substrate*, causes a harm while working for his employer, or when harm is in some kind of relation to his work.⁵⁶

Secondly, we have to check if we can apply culpability rules. When an employee causes harm to a third party (injured party) on purpose, the latter is allowed to require a compensation also from an employee.⁵⁷ In such situation, we define tortious liability by an employer and his employee as a *liability in solido*. The law always protects third party (injured party) and for this reason an injured party may require compensation either from an employee or its employer. If he/she requires it from an employer, the latter also has a legal claim (same amount as he/she paid to an injured party) against his employee.

In a case when the digital workforce causes harm to the third party, we should also first look at the strict liability rules. Here, just like in the previous case, certain prerequisites need to be met in order to treat harmful behavior committed by an employee as the behavior of an employer. A mechanical worker needs to be a part of employer's *personal substrate* and digital workforce, because as a part of the *personal substrate*, the machine causes a harm while working for his employer or when a harm is in some kind of relation to his work.

When we try to apply culpability rules we observe that a mechanical worker, due to its nature, is incapable of causing harm on purpose. Therefore, in case of digital workforces, employers will always be held strictly liable for digital workforce'

⁵⁶ Ibid., 568.

⁵⁷ Ibid., 570.

harmful (unforeseeable) actions or omissions, since they will not be able to transfer responsibility to digital workforce.

IX. Conclusion

The presence of artificial intelligence and robots in the medical profession reveals new risks regarding privacy of personal data and emphasizes legal gaps regarding liability for harm caused by robots, equipped with advanced algorithms. Moreover, their inclusion into medical procedures and impersonal treatment of its patient will lead to a decrease of moral note that is significant for sensitive doctor-patient relationships. As we have seen, robots achieved some amazing results while treating patients and it seems reasonable to expect and also to anticipate that they will be even more engaged in the future. However, as I have presented and analyzed in the previous sections it is obvious that law falls behind the technological innovations, which on one hand bring rapid improvement of medical practices and on the other they set new legal questions. I believe that we should not overlook robots' involvement into medical field where, due to their advanced abilities, they are capable of performing activities and tasks that used to be exclusively attached to the human workforce. On this point I want to conclude that legislators around the world should take into consideration the suggestion regarding the re-regulation of employment relations.

Received: September 1, 2017

Revised: November 14, 2017

Accepted: November 23, 2017