Advances in Technology Innovation, vol. 9, no. 1, 2024, pp. 65-84

Challenges and Solutions to Criminal Liability for the Actions of Robots and AI

Vladimír Smejkal[1,*], Jindřich Kodl[2]

[1]Department of Informatics, Faculty of Business and Management, Brno University of Technology, Brno, Czech Republic

[2]Authorized expert of cryptology and information systems security, Prague, Czech Republic

Received 20 April 2023; received in revised form 17 September 2023; accepted 23 September 2023

DOI: https://doi.org/10.46604/aiti.2023.12038

Abstract

Civil liability legislation is currently being developed, but little attention has been paid to the issue of criminal liability for the actions of robots. The study describes the generations of robots and points out the concerns about robots' autonomy. The more autonomy robots obtain, the greater capacity they have for self-learning, yet the more difficulty in proving the failure foreseeability when designing and whether culpability or the elements of a specific crime can be considered. In this study, the tort liability depending on the category of robots is described, and the possible solutions are analyzed. It is shown that there is no need to introduce new criminal law constructions, but to focus on the process of proof. Instead of changing the legal system, it is necessary to create the most detailed audit trail telling about the robot's actions and surroundings or to have a digital twin of the robot.

Keywords: robot, criminal liability, audit trail, criminal evidence

1. Introduction

Concerning robotization and the growing capabilities of robots, a novel phenomenon is emerging. It could be described as self-will, self-initiative, or the independent, autonomous decision-making of robotic systems, which deviates from the conservative position of non-thinking. Modern robotic systems are more sophisticated, and mere mechanisms are fully subject to the will and instructions of humans. Some robots currently find themselves in the new role of what is perhaps best described as a self-acting, i.e., autonomous, machine. This new aspect will undoubtedly be reflected in law. The authors, therefore, focus on the potential criminal liability for the unlawful consequences of robots' actions and how to prove it.

Legislation on civil liability is currently in the works. However, scarce attention has been paid to issues of criminal liability for the actions of robots. In the European Union (EU), for example, two pieces of EU legislation have been published which may impress the officials with the issue of liability for the actions of artificial intelligence (AI) in their entirety, i.e., the Artificial Intelligence Act (AI Act) and the AI Liability Directive. In reality, they only concerned with the strict liability of manufacturers for defective products, the rules for access to information, and the burden mitigation of proof concerning damage caused by AI.

Specifically, the AI Act [1] provides for (a) harmonized rules for the placement on the market and AI systems service for their use in the EU; (b) a prohibition on certain AI practices; (c) scrupulous requirements for high-risk AI systems and obligations for operators. Meanwhile, the AI Liability Directive [2] builds on the AI Act, simplifying the legal process for victims to prove someone else's fault caused the harm.

* Corresponding author. E-mail address: smejkal@znalci.cz

English language proofreader: Chih-Wei Chang

In the area of liability for damage caused by a product defect, i.e., in the area of civil liability, liability is conceived as an objective liability. If proven that the damage was caused by the injured party, or if it can be reasonably assumed considering all circumstances that the defect did not exist when placed on the market or occurred later, the obliged person has the possibility of liberation. In the area of criminal liability, it is unlikely to set out some form of strict liability in law because the basic principles of criminal liability include the criminality of the conduct, the consequence, and the causation between the conduct and the consequence, i.e., proving the culpability of a perpetrator or group of persons.

This paper does not construct AI and robots as entities equipped with criminal liability but describes the possibilities of addressing it within traditional legal systems. The authors disagree with claims that AI could prove to be an exceptional opportunity to change law and legal theory [3-4]. The most common tort cases are likely to emerge from poor programming. This will certainly be true until robots reprogram themselves. Attempts to hold manufacturers, programmers, and users on legal theories such as negligence or product liability under the criminal code predominate over the current level of "weak" AI, while the view of recognizing direct criminal liability for AI itself predominantly supports the "strong" AI in the future [5].

The difference between "strong" and "weak" AI is that "weak" AI aspires only to model the mind when solving partial and specialized problems, while "strong" AI is recognized as a universal model of human thinking in a software way, i.e., a replication of the human mind in a computer [6-7].

Authors need to prepare for this future. However, in the authors' opinion, it is neither by creating robots as fictitious persons endowed with legal responsibility, nor does it make sense given the purpose of punishment involving retribution, incapacitation, general deterrence, and rehabilitation of the offender [8], or even the death penalty [9]. Once humans have decision-making authority over robots, they can simply deactivate or dismantle the robot. Nevertheless, it does make sense for a person - natural or legal - who was responsible for the robot or found to be at fault. They see the introduction of corporate criminal liability, as it is conceived in many countries and has long historical roots, as an appropriate way forward [10-11]. These prosecuted legal persons may primarily be the manufacturer or operator of the robot with shared culpability, which is also an option [12].

In recent years, the use of robots has been rapidly expanding in various sectors. Particularly, industrial production (welding, assembly, material handling, etc.), construction and maintenance of infrastructure, healthcare (robotic surgery and diagnostics), logistics and warehousing, transport (driving vehicles, autonomous vehicles, trains, or planes), agriculture, domestic and accommodation services, and the military (military robots). The increasing numbers and use raise the likelihood of incidents related to robots or their malfunctioning. The purpose of this study is to explore whether existing criminal law tools are sufficient to successfully address issues related to criminal liability for the actions of robots and AI without the need to create new constructs.

2. The Term "Robot" and Categories of Robots

A broader perception of the term "robot" has emerged than before. In the 20th century, robots were perceived as devices performing physical work (the word "robot" was derived in 1921 by the Czech writers Čapek brothers from the word "robota" which means "work"). It was only much later that "software robots" implementing robotic process automation (RPA) emerged, which can be used for any process that is repetitive according to certain rules - from sending emails based on certain criteria to high-frequency trading (HFT). Both types of robots can then be controlled using AI. The following text categorizes only robots in the classical sense, i.e., physically intervening in the environment, with special attention to autonomous vehicles.

2.1. Definition of robot

The technical nature of the robot is constantly evolving from simple manipulators to autonomous systems, where the robot can perform the intended tasks based on the current state and sensor data without human intervention, moving in its

environment, and performing the intended tasks while comprising a control system and control system interface. The robot's actions must have either direct or indirect responses in the physical world (not necessarily by direct mechanical manipulation, but also by passing a command to another system having an interface to the physical world, e.g., water supply network control or road traffic control).

Software-only robots will be understood and classified as computer programs, which are not addressed in this paper. A robot, therefore, differs from information systems in its ability to directly interact with its environment physically. Furthermore, for a long time, the term robot (except household robots) was largely understood as an "industrial robot", even though the word "industrial," i.e., a system designed to perform simple, typically repetitive mechanical operations, may now seem too finite as robots are making inroads into significantly "non-industrial" areas, such as medicine.

Robots can be fully or partially autonomous, or they can be remotely controlled (by a program in the cloud and/or by a human), often probably both in some predictable synergy. The degree of a robot's autonomy is still being determined by the human, either a priori based on legislation or in a specific situation by the person in control, where the human operator is always part of the control loop of such a robot's program.

An example is the autopilot in an aircraft, in which, after setting certain conditions, the pilot hands over control of the aircraft, but oversees its operation and can intervene in this robot's work at any time – correct its control of the aircraft, or switch it off completely. However, the conditions under which the pilot may hand over control of the aircraft to this system are defined in the aviation regulations and the aircraft manufacturer's regulations by specific flight conditions.

2.2. Generations of robots

Robots can be divided into five generations according to the level of their ability to make autonomous decisions:

(1) The zero generation includes manipulators and robots usually without feedback, where any malfunctions or changes in the monitored areas (signaled by sensors) result in the next step being disallowed and the system being stopped (the so-called "central stop") while the maintenance staff is called.

(2) The first generation includes robots with simple feedback capable of switching between several deterministically operating subprograms (developed in advance by a human) and working.

(3) The second generation includes robots with optimization capability, which is the ability to select the optimal program from predefined programs based on a specified criterion, i.e., the precise rule governing the decision about the next known action.

(4) The third generation is characterized by robots capable of independently modifying the original program (action plan) with a posteriori knowledge. Here, only the goal of the activity (task) is predefined, while the method of achieving the goal is left to the intelligence of the control system, which itself creates an action plan consisting of successive steps and activities to achieve the given goal. An action plan is a sequence of robots starting in an environment, described either numerically or symbolically – by logical statements that the robot interprets per se to achieve the defined global goal [13]. The formulation of the plan then consists of finding a way in the state space from the current state to the target state.

(5) The fourth generation is represented by autonomous robots with social, human-like behavior, which means they choose the goals of individual tasks independently based on an appropriate global criterion, e.g., the principle of long-term existence/autonomy of such a system (survival, energy saving, etc.).

Since the third generation, determining the cause of undesirable behavior is likely to be a problem and any search for legal liability will have an unparalleled difficulty.

More or less intelligent robots are becoming useful autonomous elements in complex industrial production sections. Their ability to make independent decisions, and select the most appropriate tasks will increase, as their will upon autonomous reasoning [14]. There are alarming predictions that AI will reach human capabilities by 2029, and humans and machines will gradually converge, reaching the point called the singularity in 2045. It is often debated whether robots will eventually dominate humans to take full control, i.e., it's estimated that Kurzweil's Singularity will occur in 2045 [15]. Similar catastrophic predictions have been published by renowned figures such as Steven Hawking, Bill Gates, and Elon Musk. In January 2015, they jointly signed an open letter on AI with other AI experts, and the letter was to call for research into the social impact of AI to prevent some potential "pitfalls" of AI, which is said to have the potential to eradicate disease and poverty, but scientists must not create something that cannot be controlled. The open letter entitled "Research Priorities for Robust and Beneficial Artificial Intelligence" details the research priorities in an accompanying 12-page document [16].

Some authors insist that evolution will always be under human control. However, robots do not yet have the consciousness and self-awareness, which is a necessary precondition for taking over control and the estimated point of reaching the singularity because we are approaching 2045, continuing to be postponed in scientific works [17]. There is also the question as to whether robots will ever become self-aware because the necessary prerequisites are the natural emotions of machines, the ability to be aware of their position in the world models, legal and moral rules, power/priorities, as well as communication-based on these models with other robots in the community, especially to plan actions towards a certain goal that the robots themselves have set [18]. Unless humans enable robots to set their own goals or activate themselves, the singularity is expected not to happen at all [19]. Nonetheless, the existing third-generation robots may already pose a problem.

2.3. Self-driving cars

Some evident examples of semi-autonomous robots except autopilots are so-called autonomous (self-driving) vehicles. As the decisions are left to the robot's control system, we will be more interested in:

(a) Setting the decision algorithms so that their operation or failure does not violate life, health, or property. According to the authors, the tentatively correct solution is to set priorities corresponding to a human driver. Nevertheless, it varies especially with a cooperative strategy. Moreover, in this context, it is indispensable to mention the principle of necessity - an act that would otherwise be a criminal offense is not a criminal offense if a person commits such an act to avert an imminent threat to an interest protected by criminal law.

(b) The ability of the driver to take control of the vehicle whenever needed. In other words, a pilot can switch off or "override" the autopilot.

(c) The degree of protection of the control system against negative influences from the environment, either intentional (hacking) or unintentional (electromagnetic interference or operational failure), external mechanical influences (coming from the environment in which the vehicle is moving), and mechanical interference (by third parties), starting with maintenance work or an attack on the vehicle as a movable asset, which the control system should also be able to approach or warn at least.

(d) Self-documentation capabilities, the ability to create an audit trail of the vehicle's operation (similar to the existence of "black boxes" in aircraft), are required.

Further development is expected to lead to a cooperative strategy that will further optimize the behavior of such systems working in a group (group intelligence), which will be applied by individual autonomous vehicles among themselves – the so-called "cooperative driving." Real-time vehicle-to-vehicle (V2V) and vehicle-to-traffic infrastructure communication will improve the information available to each road user. This will make almost any active involvement of the driver unnecessary and undesirable in many cases. The point is that cooperation between vehicles, or between vehicles and their environment, can

incur the perplexity of information for the drivers to process in real-time. Automated driving will therefore be effective where the driver would no longer be able to process the massive information provided in the required time, and even following the recommendations would be insufficient. The driver will thus be de facto (not de jure) useless except the situations where the driver must be able to return the vehicle to its original, less complex mode and to take over the control if, for example, the cooperating network collapses (e.g., due to a lightning strike).

3. Liability for the Robot's Actions

To search for who is responsible for a crime committed concerning the use of a robot, we must apply the peculiarities of the robotic world to classical schemes of criminal responsibility, especially when resulting in the involvement of AI. Initially explicit, deterministic processes can change into unpredictable and non-deterministic processes. It is therefore necessary to discuss whether the traditional principles of criminal law will hold up under the new conditions.

3.1. A new dimension of liability

The new dimension of liability will be based on the presence of a relatively simple line of action and consequence concerning potential liability for perhaps all machines and equipment falling under the zero, first, and second generation that has been used hitherto. Whether it is the design or manufacture of these machines, which starts with a project where parameters are always verifiable. Therefore, the relatively simple way of subsequent checks is whether any important fact has been omitted in the design. The same applies to the manufacturing process assuming that the design is flawless on the material, the question is whether this process was followed or whether, for some reason, a deviation emerged, which may be related to the ensuing unlawful consequence. The next stage is the actual use of the final product. the utilization is regulated by law if it has complexity and the potential to harm. Typically, this applies to all means of transport where the legislation explicitly provides the means of transport can be operated on public roads.

In other cases, only liability for any unlawful consequences is regulated, such as liability for damage caused by an operational activity. In addition, many technical regulations and technical standards defining product characteristics are rendered. However, the question lies in the obligation to apply, whether the presence or absence of regulations and standards. Despite the propriety because of the duty to take preventive measures, it will be up for discussion here.

Nonetheless, it is no longer the case with third- and fourth-generation robots. Meanwhile, the regulation of the use, possible liability for damage, and harm resulting from robot failure will be substantial. In the case of a consequence contrary to the purpose and intent of the machine or device, the behavior and its compliance with the law, possible fault as a subjective element of a possible criminal offense, and the causal relationship between this behavior and the consequence could be analyzed, however, several questions will arise as: to what extent this failure could have been foreseen in the design of the robot, whether any act occurred at this stage or perhaps during the manufacture, which could be causally linked to the consequence and thus whether it is possible to establish fault or the elements of a particular criminal offense were accomplished.

Nevertheless, it will always be necessary to consider the absence of determining beyond doubt the resulting behavior under all possible circumstances in the simple case of a deterministic behavior given the random (unexpected) input signals and data (observations of the surrounding world). It creates a considerable space for technical and legal analyses focused on the foreseeability of errors in the design and programming of the robot, the prevention with appropriate modifications, and the category of *force majeure*.

3.2. Tort liability in robotics

Presumably, the existence and mass use of robots after the third generation will not change the basic paradigm. If the main or sole cause of injury is the failure of a machine, it is important from a criminal law perspective whether or not the

failure is based on fault. It is well known that civil law also recognizes a strict liability for a consequence, but that is not what is at the focus of the analysis. It is necessary to analyze what can be regarded as "fault" within the meaning of the Criminal Code, i.e., the involved fault and is the moment an event that has occurred through no fault of any individual or a legal person, for instance, a failure or an accident happened through nobody's fault but an unfortunate coincidence.

This bipolar scheme of fault vs. mischance could certainly apply to robots, insofar as we simplify the problem, disregard the actions of persons not criminally liable, and introduce actions in circumstances that preclude criminal liability (for instance, tolerable risk). But the problem lies in the distinction, which is onerous to determine a robot's autonomy and the ability to learn or self-program. Specifically, learning is simply setting the parameters of a fixed system, situation classifier, neural network, production system, etc., whereas self-programming can significantly modify the decision-making principles influencing a robot's actions.

(a) Undoubtedly, robots or robotic systems performing identically repetitive operations under controlled external conditions (typically industrial, assembly robots/manipulators) have been employed for many years. Manipulators on a production line, unless someone directly steps into their operation field, are virtually harmless concerning the arguments below.

(b) Some robots operate in environments with variable external circumstances, i.e., environments with uncertainty. These circumstances include the possibility of a collision with a human, another robot, or some other movable or immovable asset. Typically, these are robotic cars, various autonomous mobile logistics robots, etc. Nevertheless, these are predefined and predictable robots despite the broad operating range. In other words, everything operable has been programmed by humans.

(c) Autonomous systems, self-learning robots, intelligent robots, hybrid robots, etc., are complex devices. Meanwhile, they are sometimes regarded as a combination of a "living" (biological structure) and "non-living" component, whose reactions and procedures are unpredictable due to self-programming within this self-learning framework or the adoption of behavior patterns on a person or a group of persons. However, as the degree of autonomy increases, we know less about the "intracerebral" situation of such a device or the state before a certain incident.

The limit of criminal liability for all three categories is the definition of negligence in Section 16 Criminal Code of the Czech Republic, Law No. 40/2009 Coll:

(1) A crime is committed by negligence if the perpetrator (a) knew they could violate or endanger an interest protected by the Criminal Code in the manner specified in this Code but relied without reasonable grounds on not causing such violation or endangerment, or (b) was unaware their conduct may cause such violation or endangerment, although they should and could have been aware of this because of both overall and personal circumstances.

(2) A crime is committed by gross negligence if the perpetrator's attitude towards the requirement of due caution shows a manifest disregard for the interests protected by the Criminal Code.

If intent, gross negligence, and wilful negligence are excluded. It is relatively easy to determine for category (a) machines what the person operating the robot should have been known despite the alleged ignorance due to the feasibility of predicting what can be expected from the robot.

The situation is more complicated for category (b). Here, it may be difficult to establish what the person operating the robot should have known for their actions to be considered as a fault in the form of conscious negligence. For example, to what extent can the driver rely on the system controlled by active radar to stop his car in front of an obstacle when it normally does so, and this is clearly stated in the instructions? Moreover, postulating the car is equipped with a fault indicator for the system, is reliance on the system a valid reason not to have your foot on the brake when the car approaches an obstacle? Or, what is even more dangerous, when someone suddenly steps into the road?

The use of virtually autonomous robots as discussed in category (c) will pose even greater problems. For this group, it will presumably be difficult to prove intent. For example, if a "thinking" robot builds another robot instead of lifting loads and loading the person operating the robot and loads into a crusher. It will be the technical cause of the robot's failure primarily investigated. On the other hand, however, similarly to any other industrial accident, it will also be necessary to investigate whether the operator or the manufacturer (of the original robot) implemented possible prevention of such failure. Preventing a "thinking" robot from producing a defective product instead of a functional robot is a typical example, which in turn means finding out what the "thinking" robot produced and, more importantly, why.

Generally, there are two types of liability: liability under the Civil Code and liability under the Criminal Code. Meanwhile, liability for misdemeanors and administrative offenses related to robots can also be added. For criminal liability, there is the requirement of intentional fault unless the Criminal Code expressly provides sufficient negligent fault. Concerning intent, a distinction is made between direct and oblique intent, where the perpetrator is cognizant of unlawful consequences or understanding (the understanding also means the perpetrator accepts or reconciles himself to the fact that he may violate or endanger an interest protected by law as described in the criminal law).

It will be relevant whether the potential perpetrator has been obliged to behave in a certain way by law, specific regulations, or even internal rules when operating the robot. In the context of robots, these may generally be such regulations as the Labour Code in the area of occupational health and safety, the protection of public utilities, or the EU regulations on medical devices as indicated in Regulation (EU) 2017/745. Moreover, whether the person has acted intentionally or negligently in breach of their duties, which may be qualified as intention, negligence, or omission. In this context, the ruling of the Supreme Court of the Czech Republic on 12 October 2017, file ref. 6 Tdo 1062/2017-28 seems to be relevant:

"*A breach of an important duty within the meaning of Section 143*(2) *of the Criminal Code cannot be understood as a breach of any regulation, but only of such duty laid down therein, where a breach of such duty usually leads to a threat to life and health of humans, if such a consequence can easily occur and often does occur due to such breach. In assessing whether a duty is an important duty, the key criterion is to consider the consequences of breaching a particular duty and the likelihood of those consequences occurring. For an omission to amount to an act, it must be the omission of a specific duty arising from the specific position of the wrongdoer, that is a situation in which society expects and relies on the actions of the specific person. First of all, the person who has a specific duty to act must be in a specific relationship to the interest protected by law. If this specific duty to act cannot be inferred from the position of the wrongdoer, the act as a condition for criminal liability is absent.*"

It is in contrast to the consideration of reasonable or tolerable risk. Even in everyday life, we cannot avoid situations in which the occurrence of harm cannot be entirely excluded. Therefore, it is necessary to consider in each case to what extent the occurrence of a harmful consequence could have been foreseen in the first place and what measures have been or could have been taken to minimize any harmful consequences. According to Section 31(2) of the Criminal Code of the Czech Republic, *there is no tolerable risk if an activity endangers the life or health of a person without their consent with the activity being given following other legal regulations, or if the result seeks to achieve is manifestly disproportionate to the level of risk, or if the performance of the activity is contrary to the requirements of other legislation, the public interest, principles of humanity or good morals.*

The assumptions will thus be based on the statement above to ensure a robot is understood as a system. In addition to the purely mechanical components enabling the robot to do something, the robot mainly consists of computer hardware and software. Despite not being biologically alive, it is capable of self-learning through experience and interaction and autonomous thanks to sensors or data exchange with the environment. Beyond that, it can adapt its actions and activities to the environment. In other words, the robot's external behavior is determined by the primary internal configuration (the initial program developed

by humans). On the other hand, with the subsequent modification of this program by the robot, various steps were presented as parameterization, creation of a database of patterns (models, heuristics), and at a higher level by the modification of the initial neural network based on the aforementioned interactions with the environment, i.e., what is closest to the human concept of "learning."

In such a case, it may be extremely difficult to establish a causal link between the failure of the robot and the actions of manufacturers and programmers. It is only after this causal link has been established (proven) beyond any doubt that the potential fault of an individual or a legal person can be examined regarding potential criminal liability.

3.3. Negligent fault in robotics

It can be assumed that negligent crimes will predominate in robotics, although the "reprogramming" of a robot into a "killer" cannot be ruled out. Concerning criminal negligence, two basic categories of perpetrators must be distinguished. The first group includes perpetrators knowing their conduct (omission) could cause harmful consequences but relying without reasonable grounds on not causing such consequences. The second group includes perpetrators not knowing their conduct (omission) could cause harmful consequences, in which case the condition for criminal liability is the fact that they should or could have known this by considering the circumstances and personal situation (in particular their education, profession, position held, etc.).

Given the desire of perpetrators to make quick money is endless, it is possible to envisage offenses committed with gross negligence concerning robots when the perpetrator's attitude towards the requirement of due caution shows a manifest disregard for the interests protected by criminal law. Without a volition element (for criminal negligence), the law defines negligent fault through an intellectual component. Regarding conscious negligence, the perpetrator is cognizant of the possibility of the relevant criminal consequence. On the other hand, as for unconscious negligence, the perpetrator is not cognizant of this possibility. The criterion of negligence is the exercise of the required degree of caution, which could be either general (required of everyone) or specific, or higher (for instance, in the performance of certain activities or professions) – see e.g., Award of the Constitutional Court of the Czech Republic of 31 May 2016, file ref. III. ÚS 2065/15-2 or ruling of the Supreme Court of the Czech Republic of 27 June 2012, file ref. 5 Tdo 540/2012.

The term "the required degree of caution" emerges here. This is a vague legal concept with a wide range of possible interpretations. The determination of the required (minimum) degree of caution could be sought in specific regulations such as technical regulations, technical standards, or the generally accepted rules resulting from the level of knowledge in a particular field as shown in the ruling of the Supreme Court of the Czech Republic of 30 January 2014, file ref. 6 Tdo 1450/2013 concerning reasonable caution.

Based on the antecedent statements, something similar is defined as the state of the art in the Inventions and Improvements Act, which comprises everything made available to the public using a written or oral description before the date from which the applicant is entitled to claim priority. If a robot commits an unlawful act due to the outdated design, the situation is similar to a non-lege artis medical practice.

3.4. Causality in robotics

As mentioned above, it will undoubtedly be difficult to establish and appropriately prove the causal link between the behavior and consequence. In medico-legal disputes, where the initial input and the outcomes are known, the actual course is entirely unclear, and the presence of the so-called black box phenomenon has emerged [20]. Predominantly, it is about fault (usually unconscious negligence) and the predictability of the outcome. The causal link is usually undisputed unless there are multiple factors at work.

If a robot of the third and especially the fourth category (as described above) fails significantly, finding a causal chain will often be a major problem with an uncertain outcome. Here we may not always know all the factors having influenced the functioning of an autonomous mechanism. We may know its initial settings (or what they should have been if there had been no design or manufacturing error), but the question is whether we can find out all the information that influenced the robot's operation. Besides, although currently, it is still easier to do than to find out all the processes taking place in the human body, the examination of the "black box" may not be successful or even feasible under certain circumstances (in the case of a biorobot) if we consider it to be the robot's hardware and software.

Another problem may be potentially multicausal, i.e., there may be concurrent causes or accumulation of several possible causes from adversity or unlawful event, from which it will be necessary to select one cause as the decisive cause in the given case. This may be the simultaneous combination of input data (i.e., the robot's environment), the program that evaluates the data, and hardware (e.g., a mechanical hand) that performs the movement that gives rise to the unlawful act. One possibility might be to express the percentage rates of these influences if possible.

In this regard, it seems appropriate to quote the ruling of the Czech Supreme Court of 27 February 2002, file ref. 3 Tz 317/2001, according to the statement "the causal link between the act of the perpetrator and the consequence is not broken if, in addition to the act of the perpetrator, there is another fact which contributes to the occurrence of the consequence, but the act of the perpetrator constitutes the fact without which the consequence would not have occurred."

An example would be a robot's motion controller not anticipating that the robot's arm might go into a certain position simply because it is programmed to exclude that position. If someone forces the robot's arm into such a position, thereby affecting the functionality and safety of the robot in such a way that it subsequently moves the arm into a different and unpredictable position causing injury, the programmers cannot be held responsible because this effect would not have occurred in the first place without the intervention in question. Given that, the authors can conclude one of the main requirements for robots: the maximum creation of an audit trail telling about every step the robot has taken and its internal state.

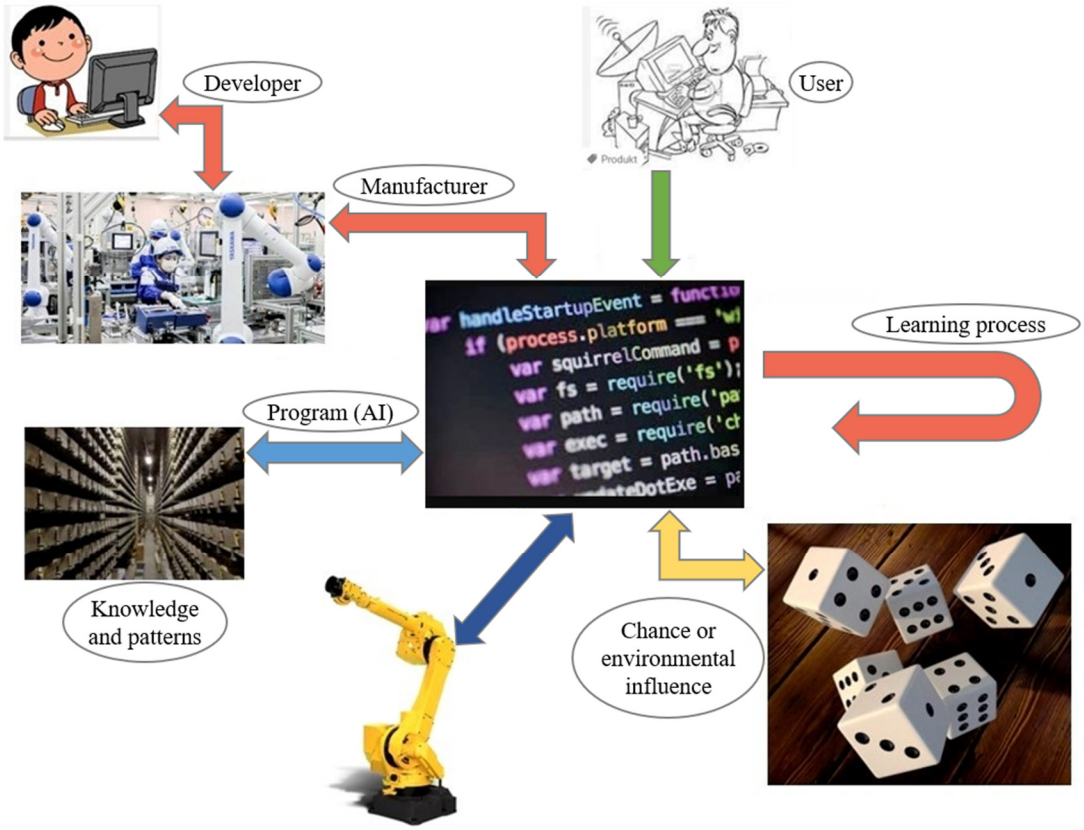

Fig. 1 Factors influencing the behavior of self-learning robots [own picture of the authors]

Fig. 1 shows all the factors that can affect the current version of the robot control program. The initial design of the robot and its setup originates with the developer and manufacturer. An empty programming tool (e.g., a neural network) learns based on the knowledge and patterns chosen by the manufacturer, but the user or the AI itself may also be involved. The behavior of the AI system is not only dependent on the outcome of the learning process but also the robot's environment (sensor information) including the behavior of other robots (e.g., in autonomous cars). The specific state of the system can therefore change expeditiously.

3.5. Robot software and tort liability

One of the main problems with robotics is the current unavailability of software with a guarantee of reliability and invulnerability. The well-known Murphy's Laws of programming state attest that "There is no program that is completely free of bugs, and that each time you remove a bug, you introduce another, hidden and more insidious bug into the program. This means that it is not possible to create a program that is completely free of bugs. Any software product will always contain some bugs."

An example leading to criminal liability could be a programming error with a certain combination of parameters (e.g., the position of the robot, its arm, load, etc.), causing an unpredictable state or behavior that can impact the environment, e.g., a mechanical attack on surrounding objects. Similar cases are known from the past in the field of space exploration, where rockets or satellites had to be destroyed due to software errors (sometimes directly in the program logic, sometimes due to a "simple" typing error).

Primarily, it seems possible to consider the liability of the author(s) of the program, who in the least "did not know that their conduct was likely to cause such an infringement or endangerment, although they should and could have known it, having regard to the circumstances and their situation." This brings us back to the question of fault.

In this respect, the Constitutional Court's ruling is relevant according to the statement "the limits of the circumstances that the perpetrator can or cannot foresee cannot be defined only hypothetically (for then everyone would have to foresee essentially everything), but must always be based on the existing objective circumstances resulting from a particular life situation, which can be characterized by a variety of factors that the perpetrator perceives with his senses and can then evaluate them according to his knowledge and other subjective dispositions. In terms of negligent fault (Section 5 of the Criminal Code), this means that, in addition to the degree of caution required by the general rules of safe conduct, there is also a subjective component, which consists of the degree of caution that the perpetrator can exercise in a particular case. There can only be fault by negligence when, at the same time, duty and the possibility of foreseeing an injury or threat to an interest protected by criminal law are present [21]." However, the question is whether it was within human power to foresee the occurrence of such coincidence (e.g., in the input data) to which the software should have been able to respond.

There may be various causes for robots' unpredictable behavior, ranging from hardware defects to software bugs. Common causes include:

(1) Hardware defects: Robots are made up of sensors, actuators, and control systems, which can fail and cause unpredictable behavior.

(2) Software bugs: Robots rely on complex software systems, and bugs in the code can cause unexpected behavior.

(3) Incorrectly calibrated sensors: If the robot's sensors are not correctly calibrated, it can lead to incorrect readings and unpredictable behavior.

(4) Environmental interference: Robots can be affected by external factors such as electromagnetic interference, which can cause unpredictable behavior.

(5) Insufficient testing: If a robot has not been thoroughly tested, it may behave unpredictably in real-world scenarios. The quality of testing is the key to minimizing the risk. It is appropriate to proceed by decomposing the system components into individual functionalities that can be described by simple finite automata whose behavior can be comprehensively analyzed to a large extent.

To address unpredictable robot behavior, it is important to diagnose the root cause through thorough testing and debugging. In some cases, hardware components may need to be replaced, and commonly the software may need to be updated or completely rewritten as well.

In this respect, it is probably appropriate to start from the principle of "required (necessary, reasonable) degree of caution," which should be embedded in the principle of "secure by design" by the robot manufacturer, which means that the designed product (which could be understood as the whole robot or just its software) is designed to be secure from the very beginning of its development and in all its stages [22]. The problem is that neither the hardware nor the software can be completely free of all errors with 100% probability. Additionally, by the requirement for reliability of communication with a system reliability of 99.9999%, the downtime of a robot can be 31.5 seconds, which is incredibly long!

Most texts dealing with safety and information systems or robots emphasize that a necessary step to increase safety and thus reduce the probability of an operational accident and possible subsequent liability is the existence of a risk management system as shown in ISO/IEC 20000 IT service management. When assessing specific cases, it will likely be a question of fact and subsequently of law, as to whether the risk was normal or extraordinary, predictable or unpredictable, and what measures were taken to minimize it.

Whether the risk was tolerable is primarily a matter for the court. However, when assessing risk in robotics, it is very likely that not only the current state of knowledge will have to be considered, but in the event of a leapfrog change in technology, which is beyond the scope of that state of knowledge, leading an entirely new situation will have to be assessed. If a risk analysis has been carried out covering all foreseeable harmful situations, if the risk of an activity causing harm has been ruled out, and if it is clear that the intended (and legal) goal cannot be achieved in any other way, then it is admissible to move on to actions involving elements of risk and consequently compliance with these conditions which should be considered when assessing the circumstances precluding the unlawfulness.

The specific examples that possibly happen will vary depending on the degree of autonomy of the robot and other circumstances. There may be an error in the program that makes the robot's actions not only unauthorized but also a threat to an interest protected by the Criminal Code. Such an error is unlikely impossible, but the error and functionality of a safety mechanism stopping the robot before the incident will be crucial. One of the criteria for assessing potential liability is whether functional testing has been performed before starting the manufacture of a robot.

3.6. Other factors affecting robots' operation

It is necessary to point out other factors affecting robots' actions. Every program works with data, while some can be entered as constants by the manufacturer, variables by the user, and collected by the robot itself through sensors. Therefore, even if the program is working perfectly, the robot may act incorrectly in an unauthorized way, and it will probably be easy to find out what has happened only in the first two cases. If a given case involves the effects of external factors whether unintentional or intentional, we will depend on the possibility of analyzing the processes inside the robot.

This is typically true of self-learning robots, which are autonomous systems that can improve their performance over time based on experience, trial, and error. Their behavior can be explained using basic algorithms and models that enable them to process data, making decisions, and learn experientially. They are sophisticated systems that can adapt and improve over time, which enables them to perform increasingly complex tasks and respond to changing circumstances in real-world environments.

One common approach to self-learning in robots is reinforcement learning, which is a type of machine learning involving an agent (robot) performing actions in the environment to maximize a reward signal. The robot learns through trial and error and takes actions possibly leading to a higher reward and improving its overall performance over time.

Another approach to self-learning in robots is unsupervised learning, which means that the robot analyses large amounts of data and discovers patterns or relationships on its own without explicit instructions or labeled data. It facilitates the robot learning about its environment and predicting future events. In both cases, the behavior of self-learning robots is determined by their algorithms and the data (datasets) on which they have been trained. The quality of their decisions and the speed at which they learn depend on factors such as the complexity of their models, the amount of data to which they have been exposed, and the quality of the algorithms used to process that data.

The behavior of these robots is determined by a combination of several factors, including:

(1) Sensors: Self-learning robots typically use various sensors to sense their environment such as cameras, microphones, and touch sensors. These sensors feed data to the robot's learning algorithms, which use this information to make decisions and perform tasks.

(2) Algorithms: The behavior of self-learning robots is determined by algorithms controlling their decision-making process. These algorithms may include artificial neural networks, reinforcement learning, evolutionary algorithms, etc.

(3) Goals and tasks: Self-learning robots are designed to meet specific goals and perform tasks. For example, a self-driving car may be designed to navigate roads safely and avoid collisions, while a service robot may be designed to help people with various tasks. The robot's behavior is customized by these goals and tasks, as well as its algorithms and sensory inputs.

(4) Experience: As the self-learning robot interacts with its environment and performs tasks, it collects data and gains experience to improve its future performance. For example, a self-driving car can learn to better navigate roads after encountering challenging driving scenarios.

In summary, the behavior of self-learning robots is a complex interplay of their sensors, algorithms, goals, tasks, and experience. A robot may also behave contrary to its intended purpose or programming due to an internal defect (hardware error, but more likely software error) or an external factor, while both may or may not be relevant concerning a third party's fault.

Unintentional influence can occur through a combination of external signals picked up by the robot's misevaluated sensors as a result of an incident. A question of assessment lies in whether the situation was both unforeseen and unforeseeable to such an extent of not having been predicted by the manufacturer. If the analysis excludes that it is not an unpredicted and unforeseeable situation and the failure of the robot emerged due to reasons per se (defect in design, program, etc.), the situation should be assumed predictable to cause partially or completely by an external influence including the intervention of another person. If a human factor is found to be such an influence, it will need to be addressed (in addition to finding the particular person) the issue of intent versus negligence. One can imagine the negligent action of someone who inadvertently or entails another activity changes the physical configuration of a robot that begins to move along a different path.

In the case of intent, this may be an external attack where the attacker attempts to influence the processes taking place within the robot enough to change its functions or functioning and subsequently cause an incident. The person who can influence the robot to act to endanger an interest protected by the Criminal Code can be anyone: the owner (user, operator, etc.) who (wilfully or unwilfully) enters incorrect data into the robot – then it will probably be about assessing possible negligent conduct in an unlawful act, but perhaps even about assessing whether the elements of the criminal offense of "Damage to a record in a computer system and on an information carrier and negligent interference with computer equipment"

have been accomplished. In the case of another third party (a hacker) who can attack the program that controls the robot (but of course also the data) to make the robot do something not under its intended purpose, the situation is clearer, i.e., this will constitute a misuse of a thing (using the robot as an object or instrument of a criminal offense).

As part of the assessment of a particular case, it will be necessary to determine which of the robot's actions can be considered the result of the negligent (or intentional) conduct of an individual. In the case of zero to second-generation robots, the robot can be perceived as a deterministic automaton (DA), where all its states are known detectable or inferable, i.e., it is clear at any point of robots' actions and rationales chronologically. For the third generation of robots, where we already admit the robot's ability to independently create a program experientially, the situation is already more complex because this gives the robot the character of a non-deterministic automaton (NDA), where it can reach a certain state through multiple paths states simultaneously.

Possibly by the theory of computer science, we will be able to convert an NDA into a DA and establish the reasons for the robot's behavior. Nonetheless, it is not universally applicable. At the outset, such a robot will always have the same initial state, but once it has been trained, it will depend on many independent factors as the internal state diverges after an incident. The original instruction set and dataset used would leave the robot as a deterministic mechanism without further learning (training) of the model. Nevertheless, the application of different behavioral patterns with different outcomes will translate into the robot's architecture differently each time, depending on the patterns and models used.

In the fourth generation, when we talk about autonomous mechanisms having their intelligence, the algorithmic complexity (variability) inside the robot will be incredibly high that it cannot be ruled out that it will be impossible to determine the exact events and rationales inside from a certain point in time. AI tools such as neural networks or fuzzy logic, evolutionary programming, and genetic programming, are moving us from a world where we can tell the causation to a probabilistic one. This is a probabilistic prediction, where no one can guarantee that a given system will evolve and behave as predicted. Then we get into the area referred to as NP-hard problems, where it will be necessary to look for approximate or partial solutions giving a satisfactory answer at least in some cases. It is a question, however, whether this approach is applicable in finding criminal evidence.

Owners (operators) of robots whose actions may endanger the safety of the public in a malfunction (due to their weight, nature of the operation, etc.) or whose operation may be labeled as "particularly dangerous," which may also be held criminally liable for any accident caused by the robot. Their tort liability for negligent conduct may be at least twofold:

(1) They have acted contrary to the manufacturer's instructions (manual), e.g., overloading the robot instead of performing proper maintenance.

(2) They failed to exercise the required degree of caution given the characteristics of the robot.

In this context, it is also possible to mention the potential liability of the robot owner in the failure to perform the mandatory software update released by the manufacturer.

Švestka and Smejkal dealt with a similar issue much earlier concerning the updates of antivirus software. "In the case of computers, there is no legal obligation for their operators to use antivirus software, let alone to update it. However, even if we were to admit this, which of course could only be done by way of law (we note, though, that this is a vision that is difficult to imagine), then the determination of the liability for damage must rest on proof beyond reasonable doubt that the owner (operator, user), or which one of them, has committed the specific misconduct. If a hacker gains access to a computer and subsequently causes damage to another party – whether directly or by using someone else's computer – it will be very difficult to prove in practice whether the fault can be attributed to the manufacturer of the operating system, to an application program purchased by the user, or whether the misuse of the computer was caused by neglecting to update a "protection" program, etc [23]."

The authors believe this is still true for a robot. However, it could happen that the obligation to update the software would be stipulated in the contract or the technical conditions for the use of the robot, i.e., in documents binding on the owner. The owner would probably be liable for the damage caused under civil liability and such conduct could also be considered negligence, perhaps even gross negligence from a criminal point of view.

3.7. Legal persons as potential perpetrators of robotic crimes

For completeness, it is also necessary to mention the potential criminal liability of legal persons. This liability arises from the fact that not only an unlawful act committed in the interests of a legal person or the course of its business activities by persons such as managers but also employees or persons in a similar position, may be attributed to the legal person. However, even if an unlawful act committed by the above persons can be attributed to a legal person, the legal person may still be released from criminal liability if he has made all the efforts that could reasonably prevent the commission of the unlawful act by its managers or employees.

This construction is also decisive in the assessment of whether the legal person is liable for the unlawful consequence of the robot's actions. Regarding the accidents caused by robots, the foremost question will be whether the liability of an individual for the accident can be established, or whether criminal liability is involved as the case may be. However, it should be noted that it may not always be a question of finding a specific individual or specific individual identified as having committed an unlawful act attributable to a legal person. Consequently, the criminal liability of the legal person is not affected.

Thus, the mechanism for establishing criminal liability of a legal person for an accident caused by a robot is more complex because it does not end with the identification of the liable individual but continues with the assessment of whether the individual acted in circumstances in which their unlawful act is attributable to the legal person. If the answer is affirmative, it only remains to consider whether the legal person has exempted himself from criminal liability by considering the efforts to prevent the commission of the unlawful act by the persons. The problem of robots was brought back to the question of the extent of potential foreseeability of a robot's reaction in a certain situation and the extent of the liability to be inferred for lack of relevant foresight. It is imaginable that the criminal liability of a legal person, a manufacturer of a third or fourth-generation robot, and the employees did not foresee the robot's harmful behavior due to the insufficient understanding of the robots' design. And the legal person, their employer, failed to address this professional deficiency. More loftily, the legal person as a manufacturer may have been technically capable of creating a monster but was unable to oversee the capability of this monster.

4. Prevention and Evidence

Actions of both robots that will act directly, i.e., causing mechanical movements, and robots that will act only based on information such as robots trading on the stock market to consequently be capable of manifesting as negative or directly contrary to the law. Two problems emerge as follows:

(1) How to prevent or at least minimize the likelihood that such actions will occur? – prevention.

(2) How to establish the reasons leading to the negative actions including the liability of a particular individual or legal person? – evidence.

The situation will be quite different for deterministic robots, where it is possible to justify any time why the robot acted as it did. If the contents of its memory are available, the computer program and data will be located despite the absent possibility to predict in advance in many cases what values the data will take. In other than deterministic processes, i.e., where the robot itself decides how to react to the stimulus, the situation will be more complex. In these cases, there will be a wide range of possibilities, from the simplest systems to the most complex self-learning ones, which will be able to modify their code based on processes that cannot be predicted to a greater or lesser extent.

4.1. Prevention

The problem referred to in point 1 above will need to be addressed by a risk analysis exploring all the possible states that the robot can get into and try to set up defense mechanisms to prevent negative actions from occurring. The higher level of autonomy of the system cooperating with other systems is a typical example for autonomous vehicles. Undoubtedly, we have reached a point where the level of protective measures may prevent the full use of the system.

Minimizing the risks associated with the robot's operation involves a multi-step process, which includes the following:

(1) Risk assessment: The first step is to assess the potential risks associated with the robot's operation. This involves identifying the types of hazards the robot may encounter with the potential consequences of these hazards. This may include simulations, testing, and other methods to evaluate the safety of the robot in different scenarios.

(2) Safety design: The design of the robot should include safety features minimizing the risk of harm to humans, animals, and the environment. For example, if the robot is designed to work near humans, it should be trained to respond appropriately to sudden movements or other unexpected events. This may include physical barriers to prevent the robot from coming into contact with people or other objects, sensors to detect potential hazards, and emergency stop buttons to quickly shut down the robot during an emergency. Robots should be operated in a controlled environment and be subject to strict operating procedures to minimize the risk of accidents.

(3) Secure communication: Robots should use secure protocols for communication, such as encrypted communication channels to prevent unauthorized access or tampering with control and data transmission.

(4) Authentication and authorization: Robots should have mechanisms to securely authenticate the identity of users and devices to control access to sensitive data and functions based on roles and permissions.

(5) Data protection: Robots should have built-in mechanisms to protect sensitive data, such as encryption and secure storage to prevent unauthorized access or theft of data.

(6) Software security: Robots should be developed using secure software development practices such as code reviews and security testing to avoid vulnerabilities and security flaws in the software.

(7) Physical security: Robots should be designed to be physically secure, e.g., by tamper-proof covers.

(8) Testing and verification: Robots should be thoroughly tested and verified to ensure the intended function and proper safety features. This may include testing the robot in different scenarios and environments to identify potential risks.

(9) Training and learning: Users should be trained and instructed on how to operate the robot safely and effectively. This may include instructions on how to use safety features and what to do in an emergency.

(10) Monitoring and maintenance: Regular monitoring and maintenance of the robot is important to ensure its continued safe and efficient operation. This may include checking the robot's sensors and safety features, updating software and firmware, and replacing worn or damaged parts.

(11) Incident reporting and response: Incidents involving robots should be reported and investigated to determine the cause and prevention of similar incidents in the future. A plan should be responded to potential incidents including procedures for emergency shutdown, evacuation, and medical assistance.

Following these steps will help minimize the risks associated with operating the robot and ensure the safety of animals, and the environment.

4.2. Evidence

Regarding point 2, this is a question of obtaining enough information to determine the behavioral rationale. It means collecting a huge amount of data is not dissimilar to collecting data in the so-called black boxes found in airplanes at least in the form of monitoring the robot's last moments of operation. However, even this may not provide a clear answer to the question of who is to blame for the robot's negative actions, especially in cases where the robot will be interacting with the environment. It is important to note that sometimes the cause of the actions may be due to a combination of factors and require a combination of the above approaches to fully understand and resolve the problem.

Typically, an autonomous vehicle is surrounded by a very heterogeneous environment consisting of other vehicles, other entities, and objects located in the environment, but it will also receive information from a variety of sources, ranging from cooperating vehicles and traffic control-related signals. However, meanwhile, these can be false signals, ranging from exterior interfering signals of unintentional origin to intentional attacks.

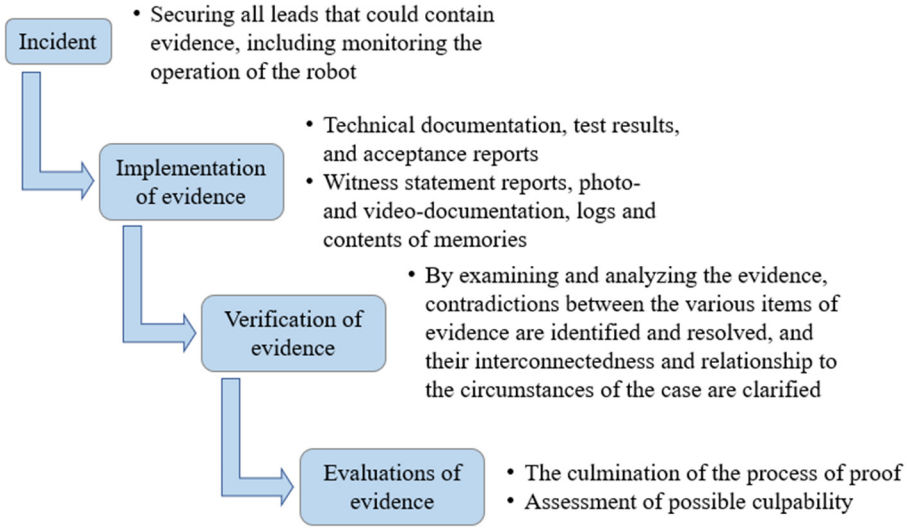

Fig. 2 Information about the incident applicable to criminal proceedings [own picture of the authors]

Fig. 2 illustrates a description of the pre-trial criminal procedure, which consists of securing clues about the robot's activities. The activities include the analysis of technical documentation, the interview with witnesses, documenting the crime scene, the examination of evidence to resolve any inconsistencies, the evaluation of both individual and collective evidence, and finally the assessment of someone's possible culpability.

4.3. Digital twins

A new phenomenon that has emerged in connection with Industry 4.0 is called "digital twins," where each physical element has its virtual representation, where its behavior and interactions with the environment are simulated by a software module. "*A digital twin is a virtual version of a physical entity, whether product, factory, or some other type of asset or system. The digital twin unites business, contextual, and sensor data to represent the physical object* [24]." Software modules, representing physical elements in a virtual space, collaboratively solve tasks, coordinate their activities, and make decisions using services they provide to each other or call up through the Internet of Services (IoS). The concept originated in products and then in machines and entire production lines [25].

According to the research report named Digital Twins in Smart Cities and Urban Modelling by analyst firm ABI Research, digital twins will become a substantial tool in city modeling and optimizing the operation of smart cities [26]. Conceivably, digital twins could be used to analyze the causes of a delinquent robot's actions, whereby individual variants could be modeled to determine the most likely course of action. Furthermore, virtualization could also be used to digitally recreate the event.

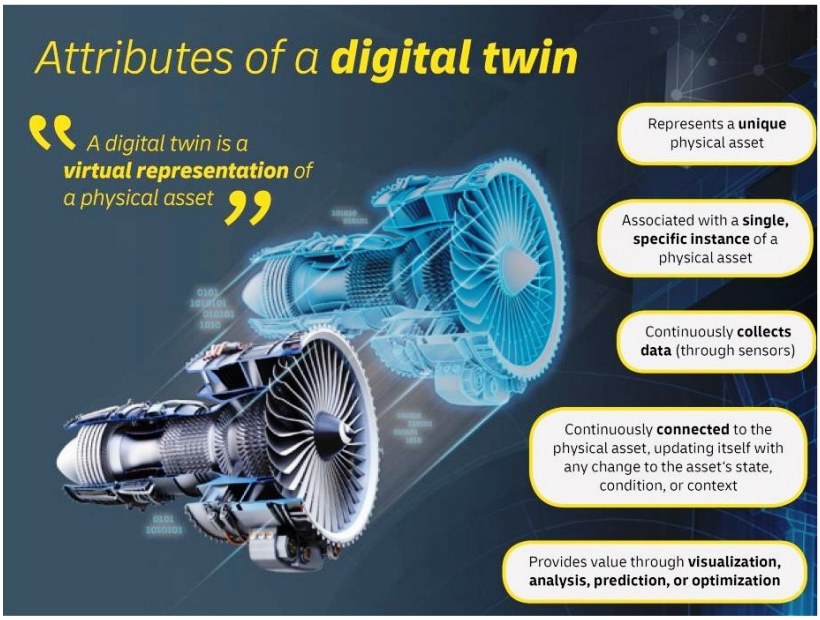

Fig. 3 Digital twins [27]

Fig. 3 illustrates the properties of a digital twin as a software representation of a physical asset. It enables us to model all possible states of the robot in the framework of prediction and analyze the behavior of the robot based on sensed variables in real operation.

In a previous article, it was written that "the first corpse doesn't count." In other words, it cannot be blamed on the manufacturer who is likely to be the main entity liable for the robot's actions due to the impossibility to foresee all states, situations, and influences. Black swans will always exist. On the other hand, if that manufacturer does not learn from all previous experience, i.e., if it fails to take into account the current state of knowledge of science and technology, then it could be liable at least for negligence.

5. Results and Discussion

From this perspective, the bottom line is that the robot either

(a) Operates in a predicted and desirable way, i.e., it operates correctly.

(b) Operates in a predicted but undesirable way, i.e., it operates incorrectly.

(c) Operates unpredictably.

In the cases referred to in (b) and (c), it is critical whether damage or non-material harm is caused in this way. If so, it is important from a criminal law perspective whether an interest protected by criminal law has been violated in the antecedent manner and, if so, who is liable for this.

In the case of the predicted undesirable operation of a robot, it is critical to know how such undesirable behavior occurred, whether it was predictable, and why measures were not taken to eliminate such operation. In the case of unpredicted robot operation, the same determinants will need to be examined, i.e., whether measures were taken to prevent or minimize the undesirable and unpredicted operation.

Generally, liability for illegal or criminal acts by robots or AI systems is a complex issue that has not yet been fully resolved by the legal system. In some cases, the manufacturer of the robot may be liable if it is proven that the robot was defectively designed or manufactured. In other cases, the owner of the robot may be liable if the robot has been used illegally such as having been programmed to commit criminal offenses. However, there may also be the factor that the programming

or data used to train the UI system, i.e., the programmer or data provider could also be held liable. Meanwhile, the liability of regulatory authorities cannot be ruled out either due to the absence of clear regulations or technical standards for the manufacture and/or safe use of the robot, while regulatory authorities may be held liable for failing to properly supervise the technology.

Ultimately, considering both general and specific circumstances, the design and capabilities, and applicable laws and regulations, liability for the actions of robots is likely to be determined on a case-by-case basis. However, as robots become more advanced and autonomous, it becomes increasingly difficult to determine who should be primarily liable for their actions.

In Article 12, the European Parliament resolution on 16 February 2017 [28] states that: "it should always be possible to supply the rationale behind any decision taken with the aid of AI that can have a substantive impact on one or more persons' lives." The European Parliament "considers that it must always be possible to reduce the AI system's computations to a form comprehensible by humans with considering that advanced robots should be equipped with a 'black box' which records data on every transaction carried out by the machine including the logic that contributed to its decisions." The problem will be whether the neural brains are artificially or biologically equipped in truly advanced robots.

As a result, we will be able to translate all the robot's decisions into a comprehensible form and the fully identifiable logic of its decisions. The black box originating from the aviation and containing all the necessary information can serve as an audit trail representing significant if not being the main or only evidence in determining the place, robot's state, and the causation of the incident. Given that, one can conclude the question posed at the beginning of this article: was it an intentional fault, negligence, accident, or a unforeseeable event – a Black Swan [29]?

There is a need for preparations in both the area of legal theory and practice reflected in the legislation and decision-making practice, the area of technology reflected in the elaboration of standards and the creation of "best practices," and in the area of criminology reflected in the emergence of new practices or sub-areas within the field of criminology. The application of methodologies and tools such as big data, AI, and digital twins could increase recall and thus help reduce entropy in specific robot-related crime investigations.

6. Conclusion

This article highlights that in the area of liability for damage caused by a defective product, i.e., civil liability, liability is conceived as a strict liability with the possibility of releasing the obliged party from liability if the damage caused by the injured party is proven or if it's reasonably assumable. Considering all the circumstances, the defect did not exist while being released or later. In the area of criminal liability, it would probably be impossible to set some form of strict liability in law since the basic principles of criminal liability include the criminality of an act, while the consequence caused and the causal link are set between the act and its consequence, i.e., proving the fault of a particular perpetrator or group of persons.

As long as AI and robots do not possess human characteristics (consciousness, subjective experience, emotions, motivation, will, creativity, social interaction, morality, and ethics), it is premature and unnecessary to create artificial legal constructs to sanction the wrongful actions of robots and AI. As the analysis shows, the appropriate solution is to apply the existing principles of criminal liability for the actions of legal persons - manufacturers, owners, and/or users of robots. To solve the problem of imputability of liability, it is necessary to focus on providing evidence in the form of the most detailed audit trail telling about the actions of the robot and its surroundings, or having digital biplanes of higher generation robots. The current state of robotics and AI does not require any other legal solution.

In the authors' opinion, the current possibilities of criminal law are sufficient to punish the unlawful actions of robots where the principles of civil law are insufficient. Inventing new social constructs consisting of punishing robots as such is impossible due to the dependence of being created by an owner, operator, or other responsible person. In the future, the

paradigm shift that other authors talk about would only make sense if robots were revolutionized as a separate entity with more power than humanity. Subsequently, the question is whether robots would create their legislation, hostile to humans. However, these considerations, in the authors' view, are not on the agenda at all. The purpose of this paper is to stress that as much as possible, and the focus should be on providing leads, collecting evidence, and implementing the principle of imputability within existing criminal law.

Future research in combating robot-related crime should combine technical, ethical, legal, and social aspects of this issue. It should also contribute to the development of effective measures to combat robot-related crime and ensure the safety of citizens. Meanwhile, four insights have emerge from the authors' perspectives. First, the behavioral identification of new threats and patterns is associated with robots within different sectors. Second, the analysis of risks and the proposal of measures are implemented to increase the technological safety of robots and autonomous systems. Third, the search for tools to analyze the behavior of self-learning robots is deployed in specific situations (incidents). Last, the investigation concerning the current possibilities of legal systems allow permitting of the implementation of criminal prosecution is deployed to propose new procedures and facts in this area. Consideration should be given to whether existing legal instruments are sufficient or whether changes are needed.

Acknowledgment

The authors would like to thank Mr. Libor Preucil, Ph. D., Head of Intelligent and Mobile Robotics & Center for Advanced Field Robotics, Czech Institute of Informatics, Robotics and Cybernetics of the Czech Technical University in Prague, for his valuable consultation.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, European Commission, COM/2021/206 final, 2021.
[2] Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive), European Commission, COM/2022/496 final, 2022.
[3] R. Calo, "Robotics and the Lessons of Cyberlaw," California Law Review, vol. 103, no. 4, pp. 513-563, June 2015.
[4] M. Simmler and N. Markwalder, "Guilty Robots? – Rethinking the Nature of Culpability and Legal Personhood in an Age of Artificial Intelligence," Criminal Law Forum, vol. 30, no. 1, pp. 1-31, March 2019.
[5] S. Ahn, "Artificial Intelligence and Criminal Liability," Korean Journal of Legal Philosophy, vol. 20, no. 2, pp. 77-122, 2017.
[6] J. R. Searle, Minds, Brains and Science, 13th ed., Cambridge, Massachusetts: Harvard University Press, 2003.
[7] M. Minsky, Computation: Finite and Infinite Machines, Englewood Cliffs, N.J.: Prentice-Hall, 1967.
[8] F. Focquaert, E. Shaw, and B. N. Waller, The Routledge Handbook of the Philosophy and Science of Punishment, London: Routledge, 2020.
[9] E. Watamura, T. Ioku, and T. Wakebe, "Justification of Sentencing Decisions: Development of a Ratio-Based Measure Tested on Child Neglect Cases," Frontiers in Psychology, vol. 12, article no. 761536 January 2022.
[10] F. Pollock, "Has the Common Law Received the Fiction Theory of Corporations?" Law Quarterly Review, vol. 27, no. 2, pp. 219-235, 1911.
[11] G. Hallevy, "I, Robot-I, Criminal-When Science Fiction Becomes Reality: Legal Liability of AI Robots Committing Criminal Offenses," Syracuse Science & Technology Law Reporter, vol. 22, pp. 1-37, 2010.
[12] L. Wang, "New Challenges Posed by Robots to China's Civil Code in the Age of Artificial Intelligence," Frontiers of Law in China, vol. 17, no. 1, pp. 33-41, March 2022.

[13] R. E. Fikes and N. J. Nilsson, "STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving," Artificial Intelligence, vol. 2, no. 3-4, pp. 189-208, 1971.
[14] N. Bostrom, Super-intelligence: Paths, Dangers, Strategies, Oxford: Oxford University Press, 2014.
[15] R. Kurzweil, The Singularity Is Near: When Humans Transcend Biology, New York: Penguin Books, 2006.
[16] S. Russell, T. Dietterivh, E. Horvitz, B. Selman, F. Rossi, D. Hassabis, et al., "Letter to the Editor: Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter," AI Magazine, vol. 36, no. 4, pp. 3-4, December 2015.
[17] M. R. Ford, Rise of the Robots: Technology and the Threat of a Jobless Future, New York: Basic Books, 2015.
[18] E. Brynjolfsson and A. McAfee, The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, New York: W. W. Norton & Company, 2014.
[19] R. Logan, "Can Computers Become Conscious, an Essential Condition for the Singularity?" Information, vol. 8, no. 4, pp. 161, December 2017.
[20] A. Doležal and T. Doležal, "Proof of Causation in Medical Malpractice Cases in the Czech Republic," The Lawyer Quarterly, vol. 5, no. 3, pp. 195-205, 2015. (In Czech)
[21] Constitutional Court of the Czech Republic, The Ruling of the Constitutional Court of the Czech Republic, May 20, 2004; similarly, for instance, Supreme Court of the Czech Republic, The Ruling of the Supreme Court of the Czech Republic, September 06, 2001.
[22] M. de la Cámara, F. J. Sáenz, J. A. Calvo-Manzano and M. Arcilla, "Security by Design Factors for Developing and Evaluating Secure Software," 10th Iberian Conference on Information Systems and Technologies, pp. 1-6, June 2015.
[23] A. Maurushat and K. Nguyen, "Correction to: The Legal Obligation to Provide Timely Security Patching and Automatic Updates," International Cybersecurity Law Review, vol. 3, no. 2, article no. 495, December 2022.
[24] L. Gould, "What Are Digital Twins and Digital Threads?" Automotive Design & Production, vol. 130, no. 2, pp. 30-32, 2018.
[25] Siemens, "Digital Twin," https://www.plm.automation.siemens.com/global/en/our-story/glossary/digital-twin/24465, November 05, 2018.
[26] ABI Research, "Digital Twins, Smart Cities, and Urban Modeling," Research Report AN-5239, September 15, 2019.
[27] Bonn, "Implementation of Digital Twins to Significantly Improve Logistics Operations," Trend Report DHL, June 27, 2019.
[28] European Union, European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), Official Journal, C 252/239, 2018.
[29] N. N. Taleb, The Black Swan: The Impact of the Highly Improbable, New York: Random House, 2008.