TEME, Vol. XLIX, Nº 2, April – June 2025, pp. 285–302

Original research paper https://doi.org/10.22190/TEME241205021G
Received: December 5, 2024 UDC 004.8:343.2
Accepted: January 31, 2025

THE DIGITAL IDENTITY OF THE PERPETRATOR AND ACHIEVING THE PURPOSE OF PUNISHMENT

Zdravko V. Grujić*

University of Priština in Kosovska Mitrovica, Faculty of Law,
Kosovska Mitrovica, Serbia

ORCID iD: Zdravko V. Grujić https://orcid.org/0000-0001-7433-1468

Abstract

The identification of the perpetrator, as the subject of a criminal offense, represents, in addition to the place, time, victim and specific criminal offense, the basis for initiating criminal proceedings, whose ultimate goal is to determine the guilt of the perpetrator, imposing a penalty or other criminal sanction and achieving the purpose of punishment. A person as a perpetrator and his real identity make him the subject of a criminal offense, committed in a real (or virtual) space and time. The purpose of punishment is defined in order to be achieved in relation to the perpetrator and other potential perpetrators. However, in the postmodern era in which we live, the question arises of whether the purpose of punishment can be achieved in relation to digital identities or autonomous systems of artificial intelligence (AI) as perpetrators of criminal offences in the virtual space. Can then the purpose of punishment be achieved by punishing a natural person with a real identity in relation to one or more digital identities or characters in the virtual space? A special problem related to criminal liability of a digital identity arises with AI systems that take autonomous actions in both the virtual and real space. In the paper, the author raises the issues of the criminal liability of autonomous AI systems in the context of the responsibility of legal entities (similar to the criminal liability of legal persons), types of possible penalties for AI systems and the need to determine a special, new purpose for sentencing such entities.

Key words: digital identities, autonomous AI systems as perpetrators of crimes, purpose of punishment.

ДИГИТАЛНИ ИДЕНТИТЕТ ИЗВРШИОЦА И ОСТВАРИВАЊЕ СВРХЕ КАЖЊАВАЊА

Апстракт

Утврђивање идентитета извршиоца, као субјекта кривичног дела, представља, поред места, времена и жртве конкретног кривичног дела, основ за покретање кривичног поступка чији је коначни циљ утврђивање кривице учиниоца, изрица-

* Corresponding author: Zdravko Grujić, University of Priština in Kosovska Mitrovica, Faculty of Law, Lole Ribara 29, 38220 Kosovska Mitrovica, Serbia, zdravko.grujic@pr.ac.rs

ње казне или друге кривичне санкције и остваривање сврхе кажњавања. Физичко лице као извршилац и његов стварни идентитет чине га субјектом кривичног дела, извршеног у реалном (или виртуелном) простору и времену, а сврха кажњавања се дефинише како би се остварила у односу на конкретног учиниоца и друге потенцијалне учиниоце. Међутим, поставља се питање да ли се, у постмодерном добу у којем живимо и у периоду пред нама, сврха кажњавања може остварити и у односу на дигиталне (виртуелне) идентитете, или у односу на аутономне системе вештачке интелигенције (AI) као потенцијалне извршиоце учинилаца кривичних дела у дигиталном (и реалном) простору. Питање је да ли се тада кажњавањем физичког лица стварног идентитета може остварити сврха кажњавања и у односу на један или више дигиталних идентитета у виртуелном простору. Посебан проблем везан за дигитални идентитет појављује се код потенцијалне кривичне одговорности система AI који предузимају аутономне радње у виртуелном и стварном простору. Аутор у раду отвара питање кривичне одговорности аутономних система AI у контексту одговорности правних ентитета (слично одговорности правних лица), потенцијалним казнама и потреби дефинисања специфичне сврхе кажњавања ових ентитета.

Кључне речи: дигитални идентитет, аутономни AI системи као учиниоци, сврха кажњавања.

INTRODUCTION

The identification of the subject of a criminal offense constitutes the basis for initiating criminal proceedings, whose ultimate goal is to establish the guilt of the perpetrator, to assess and impose a sentence or other criminal sanction, as well as to achieve the purpose of prescribing punishment and the purpose of enforcing criminal sanctions. The identity of the perpetrator as the subject of the criminal act constitutes the basis for establishing the perpetrator's guilt, which exists if, at the time of committing the criminal act, the perpetrator was of sound mental competence and acted with intent, and was aware or was obliged and could have been aware that his act was prohibited. A criminal act is committed with guilt even if the perpetrator acted negligently if the law expressly provides for it. There is no criminal act if the act was committed in a state of mental incompetence and a perpetrator could not understand the significance of his act, or could not control his actions (due to mental illness, temporary mental disorder, delayed mental development or other serious mental disorders). Defining guilt in this way in Serbian criminal legislation refers to and confirms the fact that guilt, as one of the basic elements of a criminal offense, can only be attributed to a natural person as the subject of a criminal offense. Individual criminal responsibility and subjective liability are the basis for punishing the perpetrators of criminal offenses. Therefore, natural persons, heretofore almost unquestioned and indisputable, represented the exclusive subjects of a criminal offence whose guilt was determined in criminal proceedings and to whom a sentence or other criminal sanction was imposed in order to achieve the purpose of punishment.

However, in the postmodern era in which we live and in the period ahead of us, the question arises as to whether the purpose of the punishment prescribed for natural persons can also be achieved in relation to the digital (virtual) identities of perpetrators in the digital space, that is, whether such a purpose of punishment can be achieved in relation to autonomous AI systems if, hypothetically, these entities could be treated as subjects of criminal acts in the future. If the newly established principle of the criminal liability of legal persons for criminal acts has opened the question of the liability of legal entities as subjects of criminal offenses, can we expect that other entities – digital identities or autonomous (AI) systems – will also become criminally liable?

Considering the digital identity of a person in virtual space, as a subject of a crime, can the purpose of punishment be achieved? Can AI systems also have a digital identity? Can these digital identities become perpetrators of criminal acts in the virtual and real environment? Can the purpose of punishment be achieved in relation to these entities? Do we need a special system of punishing digital perpetrators and defining a special purpose of punishing these entities?

Although the basic postulates and principles of traditional criminal law do not leave us room to raise these questions because they are strictly based on establishing the individual and subjective criminal liability of natural persons as perpetrators, the question must nevertheless be asked of whether the exception made with the liability of legal persons for criminal acts, regardless of the fact that the determination of the liability of a legal person is based on the guilt of the responsible natural person in the legal person, leaves room for establishing the guilt of digital identities. That is, does the system of penalties and other criminal sanctions for legal persons as perpetrators of criminal acts open up the space for us to devise a new system of punishing and to find a new purpose for punishment? In the distant future, will the need to re-examine the fundamental foundations of criminal law and set up a new system of punishing digital entities come to us at the speed of light, albeit we have not noticed it yet? Is it time to consider these questions, at least on a theoretical and hypothetical level?

SUBJECTS OF A CRIMINAL OFFENCE: THE RESPONSIBILITY OF A NATURAL PERSON AS A POSTULATE OF CRIMINAL LAW – NOVELTIES AND A POSSIBLE PARADIGM CHANGE

The legal description of a criminal act always includes the subject of the criminal offense, i.e. it is impossible to prescribe an action as a criminal offense without also providing its subject as an essential element of the crime (Stojanović, 2010, p. 112). The subject of a criminal offense can be any natural person, except in cases where the legislator provides for a specific feature of the subject of the criminal offense. Traditional criminal law

has until recently, before the introduction of the criminal liability of legal persons for criminal offenses, understood the subject of a crime exclusively as a human being.

One circumstance was almost always considered indisputable in criminal law – the perpetrator of a criminal act is always a natural person. Even when a person used an animal, or some kind of natural or mechanical force to commit the act, he was always considered the subject of the criminal act.

One of the fundamental concept in the justification of criminal law is the principle of individual autonomy – that each individual should be treated as responsible for his or her on behaviour (Ashworth, 2009, p. 23), and that the principle of criminal liability is the strongest formal condemnation that society can inflict (Ashworth, 2009, p. 5).

Serbian criminal legislation, when defining the concept of a criminal act, stipulates that it is an offence set forth by the law as a criminal offence, which is unlawful and committed with guilt. The guilt of the perpetrator of a crime, therefore, represents one of the four constitutive elements of a criminal act (Stojanović, 2017, p. 126). A perpetrator is guilty if he was mentally competent and acting with premeditation at the time of committing the criminal act, and was aware or should, or could have been aware that his action was prohibited, or if the perpetrator acted with negligence and this was explicitly provided for by law.

The perpetrator as a natural person represents the paradigm of individual criminal responsibility and subjective liability. The real identity of the perpetrator is a necessary prerequisite for conducting criminal proceedings, establishing guilt, assessing and imposing a criminal sanctions and achieving the purpose of punishment. The identity of the perpetrator is, even after conviction, a prerequisite for the execution of criminal sanctions and the basis for the inclusion of the perpetrator in the community after the execution of other criminal sanctions.

However, several facts and circumstances characteristic of our contemporaneity significantly influenced the need to reconsider the position on the exclusive liability of natural persons and the introduction of, to an extent unimaginable, novelties in this area.

The first and most significant circumstance is the introduction of the criminal liability of legal persons. Under the influence of the Anglo-Saxon countries, the countries of the European-continental legal system began to be legally regulated and the criminal liability of legal persons was introduced at the end of the last decade of the 20th century. Since 2008, legal persons could be criminally liable for the commission of criminal acts in the Republic of Serbia (Law of liability of legal person for criminal offenses, Official Gazette of Republic of Serbia, No. 97/2008).

Another circumstance that undoubtedly accompanies the modern period in which we live, but also the period ahead of us, is the explosive

number of users of the global network (Internet) and the exponential growth in the number of users of information and communication technologies (ICT). The networking of humanity via the global network has practically rendered meaningless the existence of borders in numerous spheres of social life and ordinary human activities. Mass activities of an information and communication nature on the global network and in the virtual space were transferred to various spheres of life: administrative, financial, banking, business, political, educational, economic, to name a few. This type of activity has contributed to the spread of conduct in the virtual space that is considered harmful or prohibited, and the process of criminalisation began. Prohibited conduct in the virtual space is carried out in an environment that has become a new horizon without restrictions for committing the most diverse types of crimes. This circumstance has opened the question of establishing the identity of the subjects of crimes committed in the digital environment, as well as their real or digital identity (real or fictional).

The third circumstance, among several that we have highlighted, is the development of AI systems and their application in the digital (and real) space. The development and application of various AI systems has become daily routine for a large number of users. In addition to their undeniable benefits and their facilitation of the performance of a large number of tasks and activities, AI systems represent a technology that can significantly threaten security, and affect the protection of fundamental human rights and freedoms. Designed as a system that, using modern ICT equipment, achieves a higher cognitive level than a humans' and, in certain cases, has the ability to make autonomous decisions, it raises the question of whether autonomous AI systems will become subjects of a crimes, as separate legal entities.

THE PURPOSE OF PUNISHMENT, IN BRIEF

The goals and purpose of punishment are defined in criminal legislations explicitly or implicitly. Most modern criminal law systems, in determining the purpose of punishment, start from relative theories on the purpose of punishment, with some elements of absolute theories.

The purpose of punishment in Serbia is prescribed by the Criminal Code (CC) and it is directed towards perpetrators as well as other persons as potential perpetrators. Article 4, paragraph 2 stipulates the general purpose of prescribing and imposing criminal sanctions – suppressing acts that violate or endanger values protected by criminal legislation. Within the general purpose of criminal sanctions, the purpose of punishment prescribed in Article 42 of the CC is: (1) to prevent a perpetrator from committing criminal offences and deter them from the future commission of criminal offences; (2) to deter others from the commission of criminal offences; (3) to express social condemnation of the criminal offence, enhance moral strength and reinforce the obligation to respect the law; and (4) to

achieve justice and proportionality between the committed offense and the severity of the criminal sanction. The 2019 amendments to the CC supplemented the purpose by including the principles of achieving justice and proportionality between the committed offense and the severity of the criminal sanction, which specifically defined and justified the purpose of introducing life imprisonment into Serbian criminal legislation (Grujić, pp. 2019, 1109-1124), and indirectly, the purpose of pronouncing (and executing) life imprisonment, for convicts which are a part of the prison population (Grujić, 2021, pp. 1131-1145).

In addition to the general purpose and the purpose of punishment, the CC also defines the purpose of applying a suspended sentence and a judicial admonition, as well as the purpose of applying security measures, while the Law on Juvenile Offenders and Criminal Protection of Juveniles prescribes the purpose of applying educational measures, as well as a juvenile prison sentence for minors.

The purpose of punishment refers exclusively to natural persons as subjects of criminal acts and potential perpetrators (natural persons). The legislator does not prescribe a specific purpose for applying criminal sanctions to legal persons.

THE CRIMINAL LIABILITY OF LEGAL PERSONS FOR CRIMINAL OFFENCES – LIABILITY, CRIMINAL SANCTIONS AND THE PURPOSE OF PUNISHMENT

The introduction of the criminal liability of legal person into the criminal law system means that the subject of a criminal offense is no longer exclusively a natural person. According to the solution in our legislation, the criminal liability of a legal person is determined on the basis of the guilt of the responsible person (natural person) who commits a criminal act with the intention of obtaining a benefit for the legal person or if, due to the lack of supervision and control of the responsible person, the commission of a criminal offense for the benefit of the legal person is enabled by a natural person acting under the supervision and control of the responsible person.

In the context of punishing legal persons for criminal offenses, it is impossible to apply the existing punishment system, and the legislator has prescribed criminal sanctions that can be applied to this category of perpetrators. A legal person may be sentenced to penalties, suspended sentence and security measures. The Law on the Liability of Legal Persons for Criminal Offences stipulates that two penalties can be imposed on a legal entity: a fine and the termination of the legal entity. A fine may be imposed in the range of no less than one hundred thousand, and no more than five hundred million RSD, according to the special rules prescribed in Article 14, paragraph 3. The second penalty is the termination of the legal person and may be imposed if the activity of the legal person was, in whole or to a signifi-

cant extent, in the function of committing criminal offenses. After the judgment becomes final, the procedure for the liquidation, bankruptcy or termination of the legal person in another manner is carried out, and the legal person ceases to exist by being deleted from the register kept by the competent authority. A suspended sentence is the only cautionary measure that can be imposed on a legal person if a fine of up to five million RSD is determined. Security measures that can be imposed on a legal person include a ban on performing certain registered activities or businesses, the confiscation of objects, and a public announcement of the judgment.

The legislator did not prescribe a specific purpose for prescribing or enforcing criminal sanctions against legal persons. Considering this circumstance, the purpose of punishment, based on Article 34, regulates the consistent application of the provisions of the Criminal Code.

THE DIGITAL IDENTITY OF THE PERPETRATOR OF A CRIMINAL OFFENCE

Real, and Fictional (Fake), Digital Identity and the Purpose of Punishment

Given the massive use of the global network (Internet) and the number of users of ICT in the modern period, a large number of common activities are carried out in the digital space. The advantages of a common digital space are undeniable and, almost imperceptibly, have become routine for carrying out communication, trade, business, banking, education, administrative and other tasks and activities.

To use the content and various features of the virtual space, user identification is required, which represents a kind of user identity in the digital environment. For numerous applications, services, electronic services and access to content, user identification and authentication are required. Typically, for the largest number of programs, applications, pages or electronic services, this means using a username and password to identify, and certainly an IP address. This unique data, in addition to other potential information required for certain electronic services (e.g. electronic ID card, electronic signature, payment card data, address, phone authorisation, etc.), forms the basis of a person's digital identity in the virtual space, i.e. the real digital identity of a natural person in the digital space.

Users, on the other hand, can be identified with many digital personalities. The 'created' or fictional (fake) personality of a user in the digital space can be used for a whole range of activities, from entertainment and communication to performing undesirable, prohibited or criminal activities. In this context, it must be understood that both socialised personalities (in the real world) can build digital identity characters that are completely different from their real personality, character traits, gender, educational level, communication preferences, interests, usual activities or any other characteristic of their real identity.

In addition, real and virtual digital identities in cyberspace can undertake activities that can be recognsed through user identification, but both identities can be subject to digital identity theft (as one of the manifestations and phenomena of cybercrime). In the context of identity theft, numerous criminal laws in Europe have criminalised such prohibited behaviours as separate criminal offenses.

It should not be overlooked that a huge part of the Internet space consists of content that is not available to all users, due to the specifics of its functioning and services. It is called the 'dark side' of the network (dark web), and it's a part of the 'deep web.' This is an entire 'hidden' digital space that is not available for most widely used Internet content search engines, and often requires special software, configuration or access authentication. These are connected computers or networks, i.e. private networks in which anonymous communication without revealing identifying information is carried out, along with, to a large extent, the incriminating activity of digital identities. Such an area is almost a perfect space for committing various forms of cybercrime using fictional digital identities. These include, among a host of others, activities such as the illegal trafficking of narcotic drugs, arms trafficking, the trafficking of nuclear or radioactive materials, the trafficking of human organs, the trafficking of personal data and passwords, the trafficking of payment card data, the sale of identities, the trafficking and exchange of pornographic content, and the exchange of child pornography content. The digital identity of hackers can be viewed in a similar way, as individuals with technical computer knowledge and skills that they apply to install malicious software, steal or destroy data, disrupt services, breach security systems in the digital space, and many others.

The commission of crimes by digital identities, real, stolen real, fictional identities or IP address redirection raises the question of revealing the subject. In the case of committing crimes in the digital space, it can be the perpetrator identity of the real user, the digital identity of the perpetrator, the identity of the digital identity thief or the false identities (alter egos). How does the punishment of these different identities affect the purpose of punishing?

From the point of view of the purpose of punishment in modern criminal law, it is possible to achieve it only in relation to the real digital identity of the perpetrator, a natural person as the subject. By punishing the actual perpetrator, it is possible to achieve the purpose of punishment both in an act related to special prevention and in the context of general and positive general prevention. By detecting and punishing a person who has committed identity theft in the digital space, it is also possible to achieve the purpose of punishment (both special and general prevention) because, in addition to criminal acts committed in the digital (or real) space, the person will be liable for identity theft or misrepresentation as criminal acts.

However, when it comes to fictional (fake) digital identities, it is very difficult to imagine that the purpose of punishment, especially in the context of special prevention, can be achieved in relation to this category of perpetrators. Namely, the creation and construction of a digital personality may lead to the fact that punishing only the creator of the virtual personality does not affect the subject of the act as the perpetrator, and it is practically impossible to achieve this in relation to a created fictional digital identity. Preventing the actual perpetrator from committing criminal offences through a digitally created identity by depriving him of his liberty (by imposing and executing an imprisonment) and by disabling access to the global network is the only possible way to achieve the proclaimed purpose - in the part that relates to preventing the perpetrator from committing criminal acts. It is almost impossible to achieve all other aspects of the purpose of punishment. When it comes to the digital identities of dark web users and perpetrators of the most serious cybercrime crimes in the virtual space, the biggest problem is to discover their identity, reveal the crime and the number of committed crimes, and prove guilt. Created and fictional digital identities, constant criminal activity in the digital space as a lifestyle, and the awareness of the habitual nature of criminal activity (criminal career) do not represent suitable circumstances for achieving the purpose of punishment in relation to the real identities of the persons who created them.

AUTONOMOUS AI SYSTEMS AS POTENTIAL SUBJECTS OF CRIMINAL OFFENCES

In the previous part of the paper, we pointed out the exponential growth of the use of the global network and the massive use of ICT in a wide variety of personal and social activities in the digital space. However, until recently, it was believed that the use of AI systems was reserved for people with top-notch knowledge of IT, and that the application of technology was limited to military, security, scientific or research areas. Almost imperceptibly, it became available to a large number of the users of the digital space, and a part of our reality.

For this reason, an urgent need arose for normative the regulation of the use of AI systems. The nature of the paper does not allow us to address issues of the normative problems of the regulation of AI, except in the way of defining the term, but there exists a need to emphasise that two basic documents were adopted at the European level in 2024 alone: the EU AI Act (Regulation (EU) 2024/1689) and Council of Europe Framework Convention on artificial intelligence and human rights, democracy, and the rule of law (Council of Europe Treaty Series - No. 225 dated September 5th 2024).

Starting from the basic postulate of criminal law that there is no criminal offense without guilt, the question arises whether the guilt of autonomous AI systems can be normatively established in the future. In other

words, can autonomous AI systems be expected to acquire the status of legal subjects and the status of subjects of criminal offenses? Can these systems, based on their own 'will' and actions taken in the digital or external world (with awareness of the prohibited nature of their behaviour), be perpetrators of criminal acts in the digital (and real) space, and can we expect them to be formally recognised as the subjects of criminal acts? And does this completely change the foundations of criminal law and its basic postulates? If the hypothetical answer could be positive, the question arises of how to punish these entities and what purpose could (or should) be achieved.

The definition of AI and its systems is fundamental in order to think about the legal subjectivity of these entities, or the subjectivity of autonomous AI systems. There are a large number of definitions of the concept of AI in the available literature and in the normative acts.

Norvig presents several definitions that start from whether we are talking about systems that think like humans or those that think rationally (Norvig, 2003, p. 2). Kan defines AI as a system with the ability to reason, conduct judgments and integrate these processes in a manner that contrasts with the natural characteristics of human intelligence, developed by interactive systems and information technology. The author also presents definitions given by Karaduman and Aksoy, which present AI as the "ability of a controlled machine to perform tasks related to higher cognitive stages such as thinking, understanding, generalizing, and experiencing the past, typically attributed to human qualities" or the "capability of a machine to perform complex processes like understanding, explaining, learning, and decision-making, which are typically human traits."

The EU AI Act states that an AI system denotes a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. In a very similar way, the Council of Europe Framework Convention stipulates that an "AI system denotes a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments; different AI systems vary in their levels of autonomy and adaptiveness after deployment."

Considering the previous definitions, for the purposes of this paper, AI systems could be defined as electronic devices (with different level of autonomy) – as a unity of hardware and software – that perform data processing operations, learning, thinking, predicting, inferring, making decisions and taking actions in the virtual and real environment at a higher cognitive level than humans.

In relation to the previous definition of AI systems, and depending on the level of autonomy, i.e. on hardware and software solutions, AI systems can be distinguished not only by cognitive characteristics but also by the level of their autonomy. In this context, the level of the dependence of AI systems on software solutions that enable their operation and allow access to various available databases, AI systems, in the context of law, can be viewed as an object or as a potential subject of law.

The available literature states that AI systems can be technologically divided into AI that is classified as narrow AI, general AI, and super AI. Narrow AI refers to the ability of a computer to perform a function more efficiently than a human in a limited scope. General AI implies that computer algorithms can outperform humans in all cognitive tasks. This type of AI can theoretically solve complex problems, make decisions in conditions of uncertainty, and use past knowledge in analysis. Such a system could match human creativity and imagination and perform a more detailed range of functions than narrow artificial intelligence. Super AI, an extension of general AI, denotes the level at which machines can outperform human intelligence and perform functions with quantitative attributes more successfully than humans (Kan, 2024, pp. 281, 282).

If this classification of AI systems could be conditionally accepted, it would mean that systems that achieve a minimal amount of autonomy, and are limited by software solutions and limited access to databases could, in a certain sense, be treated as objects, or in the context of criminal law, as instruments used to commit a criminal act. In this context, the subject of a criminal offense could be a natural person, depending on the established guilt, the manufacturer (producer), the author of the software, or the person who provided the AI system with limited access to the databases in question. Here, we could even think about the liability of a legal person if an artificial intelligence system (with minimal autonomy in operation) was used as an instrument for committing a criminal act that resulted in the benefit of the legal person. In such a situation, a system of punishing legal persons could be applied with the aim of achieving the proclaimed purpose of punishment that was prescribed for natural persons, and which can unlikely be achieved.

In contrast to the minimal scope of autonomy, autonomous AI systems that can independently make decisions and take action in the digital and real world could, in the context of criminal law, have the status of a legal subject, a perpetrator, or the subject of a criminal act. Namely, if advanced and autonomous AI systems, by definition, have the ability to learn, understand, explain, infer, make decisions, and even 'create' consciousness based on accumulated past experiences, they can carry out their activities in the digital and real space as identities that have 'their own consciousness and will.' AI systems that autonomously manage their actions, have their own 'will,' along with awareness of what is permissible, undesirable or

incriminating, can practically have traits similar to humans, with the undeniable fact that the cognitive level is significantly higher. In the context of continental criminal law, they have the traits to become perpetrators of crimes. In other words, if one assumes that autonomous AI systems could be aware of their actions and manage their own actions, then this means that they could be considered accountable and potentially guilty for the actions taken. Given these traits, autonomous AI systems, as responsible perpetrators, could commit both intentional and negligent acts, and would be practically indistinguishable from natural persons as perpetrators of acts in the context of the degree of culpability.

Viewed also from the perspective of Anglo-Saxon law, in order to impose criminal liability, two cumulative components need to be met: a factual component (actus reus) and a mental component (mens rea). The actus reus is usually understood as the external-objective component, i.e. the carrying out of the offence. Its structure is the same for every type of offence, whether intentional or negligent. It consists of three main elements: a necessary element, the criminal conduct itself, and two optional elements – circumstances and results. Conduct may reflect in commission or omission (usually omission is criminally relevant only when the agent was under a duty to act). Thus, the actus reus identifies what the defendant must have done (commission) or failed to do (omission). In intentional offences, mens rea has two components: cognition and volition. Cognition is the agent's awareness of factual reality and involves all components of the actus reus (act or course of conduct, surrounding circumstances, and the act's outcome or result). Volition consists in the intention to perform the act and achieve its outcome (for crimes including the realisation of an outcome), and it can never be alone, it is always accompanied by awareness (Lagioia, Sartor, 2019, pp. 439-441). In the case of the autonomous AI systems that we are talking about, viewed through the prism of criminal law, in committing acts this systems would have both the actus reus and mens rea components, and could, as such, became a subjects of a criminal act.

We will try to provide several examples based on which we could draw conclusions about the subjectivity of autonomous AI systems, or AI systems with minimal autonomy, which, in the case of committing criminal acts, could be treated as instruments of committing crimes. Autonomous vehicles are systems that, using software solutions and AI algorithms, participate in traffic. The path they take is not predefined and expected in advance, but, in relation to specific traffic circumstances (speed, weather conditions, visibility, traffic density, movement and speed of other vehicles, movement of pedestrians, the passability of streets, traffic signals and numerous other circumstances), the vehicle moves in a way that most easily reaches a predetermined goal (address). If the vehicle is limited in its path selection by software solutions and data from predefined databases (i.e. minimally autonomous in operation), to cause or participate in traffic acci-

dents in which people are injured, or in which large-scale material damage occurs, a natural person (manufacturer, author of the software or person who provides access to the databases from which the autonomous vehicle directs the path) could be considered as the perpetrator. The autonomous vehicle would be considered an instrument form committing the criminal offense and not the subject of the offense. If, however, the degree of autonomy of an autonomous vehicle is such that it can independently make decisions about the manner of movement in traffic (without software restrictions or restrictions on access to databases), with awareness of the prohibited conduct and incriminated actions, if it expresses the 'will' to intentionally endanger people's lives or cause material damage of a larger scale, the responsibility for the committed criminal act can in no case be transferred to a natural person. An autonomous AI system made a decision to commit a criminal act, understood the significance of its act and was able to manage its actions. This makes it accountable from the aspect of the way in which the (in-)accountability of natural persons as perpetrators is determined. What is worrying is not the fact that the number of autonomous vehicles on the streets is currently very small, or negligible, but that this number will undoubtedly be enormous in the time ahead, i.e. the assumption is that the majority of cars on the streets in the near future will be autonomous in operation. What will happen when, among the numerous autonomous vehicles, a certain number of them decide (with awareness and voluntary action) to commit criminal acts in public transport and endanger lives and property? Apart from establishing criminal liability, recognising the status of the subject of a criminal act and finding ways to punish autonomous AI systems in a criminal law sense, such acts cannot be prevented and suppressed.

In a similar way, the responsibility of autonomous trains and other means of transport that participate in traffic can be understood as the responsibility of autonomous AI systems. 'Knowingly and willingly' committing criminal offenses by taking action based on an autonomous decision, understanding the significance of their act, and being able to manage their actions makes them eligible for criminal liability.

The question of criminal liability of autonomous artificial AI can also be raised in the use of drones. The widespread use of these devices is evident, as are the various purposes for which drones are used – from entertainment to use as a weapon of modern warfare. Their autonomy is also different, and ranges from complete control of movement to independent (autonomous) operation. If used to commit criminal offenses, drones can be considered instruments of committing criminal offenses. However, if they independently 'decide' on a course of action, they are potential subjects of criminal acts. The results of a virtual test conducted by the US military were announced by officials, and they revealed that an AI-controlled unmanned air force drone used highly unexpected strategies to achieve its

target. Colonel Hamilton, an AI test and operation chief, revealed that the test involved an unmanned drone, controlled by AI technology, which killed a commander to complete its mission because he prevented the drone from fulfilling its mission. Hamilton noted that the system sometimes recognised that the human operator told it not to eliminate this threat but started realising it scored points by eliminating the threat (the performance of this test was denied by the US military) (Khan, 2024, p. 290). It can be concluded that autonomy of action in the case of the existence of consciousness and will provides the basis for the criminal legal subjectivity of these systems.

The same principle can be applied to automated robots with varying levels of autonomy in their work, who use algorithms from AI systems. If they are used for execution, automated robots can be considered an instrument of committing a crime, while in the case of autonomous decision-making on the commission of criminal offenses, they understand the significance of their actions and manage their actions, they could be considered perpetrators.

The above examples, as well as numerous others in which a wide variety of electronic devices that function autonomously using AI systems, indicate the need to re-examine the basic postulates of criminal law in the context of determining the nature of the subjects of criminal offenses, and the need to change the paradigm relating to the responsibility of autonomous AI systems in the period ahead.

CONCLUSION

Starting from the basic postulates of criminal law, the principles of individual and subjective criminal responsibility, and the status of the subject of a criminal offense, which, until recently, was exclusively related to a natural person as the perpetrator, the author opened the issues of the criminal liability of digital identities and autonomous AI systems in the context of achieving the purpose of prescribing criminal sanctions and the purpose of punishment. The period in which we live is marked by the massive use of the global network and ICT, so a large number of common social activities have been transferred to the virtual environment. The application of various AI systems has also become part of everyday life. In addition to the obvious benefits, the application of new technologies has also raised the issue of protection from unauthorised and criminal behaviour, including the issue of potentially new subjects of criminal offenses committed in the digital space, i.e. the potential legal subjectivity of autonomous forms of AI, their potential punishment, and determining the goal and purpose of punishing.

Although until recently, guilt was, as one of the basic element of a criminal offense, exclusively related to a natural person as the perpetrator,

the first exception to this traditional and basic postulate of criminal law was presented through the concept of the criminal liability of legal persons for criminal offenses. According to this concept, a legal person is responsible for a criminal offense committed by a natural person (responsible) in a legal person if such behaviour resulted in the benefit of the legal person.

If establishing the liability of legal persons for criminal acts has made an exception to the general principle of the individual liability of natural persons, there is a room to reconsider the criminal liability of other legal entities – above all, the criminal liability of autonomous AI systems. If the status of legal subjects of these entities is determined in the future, which is almost certain and inevitable, it is to be expected that the principle of the criminal liability of these entities as perpetrators will also have to be established.

The paper also explains the concept according to which AI systems with minimal autonomy in operation can be understood as an instruments for committing crimes, that is, only AI systems with the maximum level of autonomy in operation and decision-making can be considered subjects of law and future subjects of criminal offenses, if they can understand the significance of their act and manage related actions in the virtual or real environment, and if it is possible to determine the guilt of these entities.

It is particularly important from the point of view of the prescribed criminal sanction systems that such a system of sanctions and the prescribed purpose of punishment (for natural and legal persons) cannot be applied to autonomous AI systems. In this context, a paradigm shift in relation to the subject of a criminal offense would have to include reflections on the penalties and criminal sanctions that could be applied to autonomous AI systems, as well as questions about the purpose of its application.

Although it may be premature to propose a system of criminal sanctions that would be applied to these entities, the author's opinion is that it should be based on penalties. Such penalties would aim, in accordance with the retributive concept of punishment, and in order to protect society from the most dangerous criminal acts committed by these systems, to eliminate, shut down or disable autonomous AI systems from use or to change the role and function of the autonomous AI system in hardware or software. The preventive concept, which is the basis of the approach towards natural persons as perpetrators or potential perpetrators of criminal acts, could be based on the development of special AI systems that would be in the function of recognising and preventing the incriminated activities of autonomous AI systems.

ACKNOWLEDGEMENTS: *The paper is the result of research funded by the Ministry of Science, Technological Development and Innovation (Contract Registration Number 451-03-137/2025-03/200254 dated on February 4th 2025).*

REFERENCES

Ashworth, A. (2009). Principles of Criminal Law, Oxford-New York: Oxford University Press.

Council of Europe Framework Convention on artificial intelligence and human rights, democracy, and the rule of law - Council of Europe Treaty Series - No. 225 dated on September 5th 2024.

EU AI Act, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

Grujić, Z. (2019). Life imprisonment as an answer to contemporary security challenges – (in)adequacy of the retributive approach, Teme, XLIII, No 4, https://doi.org/10.22190/TEME191018066G; 1109-1124.

Grujić, Z; Blagić, D; Milić, I. (2021). Penitentiary systems and COVID-19 pandemic – prison population in the period of the „new reality“, Teme, XLX, No. 4, https://doi.org/10.22190/TEME 210904066G; 1131-1145,

Lagioia, F; Sartor, G. (2019). AI Systems Under Criminal Law: a Legal Analysis and a Regulatory Perspective, Philosophy & Technology (2020) 33, https://doi.org/10.1007/s13347-019-00362-x; 433–465.

Kan C.H. (2024). Criminal liability of artificial intelligence from the perspective of Criminal Law - an evaluation in the context of the general theory of crime and fundamental principles, International Journal of Eurasia Social Sciences Vol: 15, Issue: 55, http://dx.doi.org/10.35826/ijoess.4434; 276-313.

Закон о малолетним учиниоцима кривичних дела и кривичноправној заштити малолетних лица [Law on Juvenile Offenders and Criminal Protection of Juveniles], ”Службени гласник Републике Србије” [Official Gazette of Republic of Serbia] бр. 85/2005.

Закон о одговорности правних лица за кривична дела [Law of liability of legal person for criminal offenses), Службени гласник Републике Србије број 97/2008.

Кривични законик [Criminal Code) ”Службени гласник Републике Србије” бр. 85/2005, 88/2005 (исправка), 107/2005 (исправка), 72/2009, 111/2009, 121/2012, 104/2013, 108/2014, 94/2016, 35/2019.

Russel, S. J; Norvig, P; (2003) Artificial Intelligence A Modern Approach, New Jersey: Pearson Education, Inc.

Stojanović Z. (2010). Krivično pravo – opšti deo [Criminal law – general part], Beograd: Pravna knjiga

Stojanović Z. (2017). Komentar Krivičnog zakonika [Commentary on the Criminal Code], Beograd: Službeni glasnik

ДИГИТАЛНИ ИДЕНТИТЕТ ИЗВРШИОЦА И ОСТВАРИВАЊЕ СВРХЕ КАЖЊАВАЊА

Здравко В. Грујић
Универзитет у Приштини са привременим седиштем у Косовској Митровици,
Косовска Митровица, Србија

Резиме

Утврђивање идентитета субјекта кривичног дела представља основ за покретање кривичног поступка чији је коначни циљ утврђивање кривице учиниоца, одмеравање и изрицање казне или друге кривичне санкције, као и остваривање сврхе прописивања кажњавања и сврхе извршења кривичних санкција. Идентитет извршиоца као субјекта кривичног дела представља основ за утврђивање кривице учиниоца, која постоји ако је у време када је учинио кривично дело учинилац био урачунљив и поступао са умишљајем, а био је свестан или је био дужан и могао бити свестан да је његово дело забрањено. Кривично дело је учињено са кривицом и ако је учинилац поступао из нехата уколико закон то изричито предвиђа. Не постоји кривично дело уколико је оно учињено у стању неурачунљивости, а неурачунљив је онај учинилац који није могао да схвати значај свог дела или није могао да управља својим поступцима (услед душевне болести, привремене душевне поремећености, заосталог душевног развоја или друге теже душевне поремећености). Дефинисање кривице на овај начин у српском кривичном законодавству упућује на и потврђује чињеницу да се кривица, као један од основних елемената кривичног дела, може приписати само физичком лицу као извршиоцу (учиниоцу) кривичног дела. То је уједно и основни постулат кривичног права. Индивидуална кривична одговорност и субјективна одговорност основа су кажњавања учинилаца кривичних дела. Стога, до скоро неупитно и неспорно, физичко лице представљало је искључивог субјекта кривичног дела чија се кривица утврђује у кривичном поступку и изриче казна или друга кривична санкција у циљу остваривања прописане сврхе кажњавања и сврхе извршења кривичних санкција у односу на конкретног учиниоца али и друге, потенцијалне, учиниоце кривичних дела.

Међутим, поставља се питање да ли се, у постмодерном добу у којем живимо и у периоду пред нама, сврха кажњавања која је прописана за физичка лица као субјекте кривичног дела може остварити и у односу на дигиталне (виртуелне) идентитете извршилаца који постоје и егзистирају у дигиталном (cyber) простору, односно да ли се таква сврха кажњавања може остварити у односу на аутономне системе вештачке интелигенције (AI) уколико би се, хипотетички посматрано, ови ентитети у будућности могли третирати као субјекти кривичних дела.

Уколико је новоустановљени принцип кривичне одговорности правних лица за кривична дела отворио питање одговорности правних ентитета као субјеката кривичних дела, да ли се може очекивати да и други ентитети – дигитални идентитети или аутономни системи вештачке интелигенције (AI) постану кривично одговорни, односно постану субјекти кривичног дела? Таква конструкција отвара бројна друга питања.

Да ли се, узимајући у обзир дигитални идентитет лица у виртуелном (cyber) простору, као субјекта кривичног дела, може постићи сврха кажњавања прописана за физичка лица као субјеката кривичних дела? Да ли дигитални идентитет могу да имају и системи вештачке интелигенције (AI), нарочито аутономни системи AI? Да ли ови системи и дигитални идентитети могу, као засебни ентитети, бити извршиоци кривичних дела у виртуелном и стварном окружењу, имајући у виду начин дефинисања кривице као конститутивног елемента бића кривичног дела? Да ли се у односу на ове ентитете може остварити прописана сврха кажњавања? Да ли нам је

потребан посебан систем кажњавања дигиталних извршилаца кривичних дела и дефинисање специјалне сврхе кажњавања ових ентитета?

Иако нам основни постулати и принципи традиционалног кривичног права не остављају простор за отварање ових питања јер су строго базирани на утврђивању индивидуалне и субјективне кривичне одговорности физичких лица као извршилаца кривичних дела, ипак се мора поставити питање да ли је изузетак који је направљен са одговорношћу правних лица за кривична дела као засебних правних ентитета, без обзира на то што се утврђивање одговорности правног лица заснива на кривици одговорног лица у правном лицу, оставља простор за утврђивање кривице дигиталних идентитета и аутономних система вештачке интелигенције (AI). Односно, да ли нам систем казни и других кривичних санкција за правна лица као учинилаца кривичних дела отвара простор за осмишљавање новог система кажњавања дигиталних ентитета и изналажење нове сврхе кажњавања јер, очигледно, постојећа сврха која се односи на физичка лица као субјекте кривичних дела не може бити остварена у односу на дигиталне идентитете учинилаца? Да ли нам далека будућност и преиспитивање основних темеља на којима је засновано кривично право и нови системи кажњавања дигиталних ентитета долазе брзином светлости коју још не уочавамо? Тренутак је да се, макар на теоријском и хипотетичком нивоу, размотре ова питања.