



Realizing the Existence of AI and the Dilemma in Criminal Law Reform: Based on René Descartes'

Marthsian Yeksi Anakotta¹, Aprikiansa Andi' Padang²

¹ Faculty of Law, Satya Wacana Christian University, Indonesia. E-mail: marthsian.anakotta@uksw.edu

² Faculty of Law, Satya Wacana Christian University, Indonesia. E-mail: aprikiansaandi16@gmail.com

ARTICLE INFO

Keywords:

AI; Criminal Law Reform; René Descartes.

ABSTRACT

This article aims to review the issue of realizing the existence of artificial intelligence (AI) and the dilemma in criminal law reform. I am of the view that realizing the existence of AI is a dilemma in criminal law reform for the future. So this article tries to explain the important role of humans in realizing the existence of AI, which can influence considerations and decision-making in making criminal law policies. The creation of criminal law is related to three basic problems: namely, criminal acts, criminal liability, and punishment. This problem is studied using conceptual approach and secondary data obtained from various literature in the fields of criminal law, Rene Descartes' thought on ethics, and the development of artificial intelligence. The result is that currently, the existence of AI can be realized as an entity that has a body without a soul (spirit). Such existence presents a dilemma in criminal law reform, the relationship between AI and the concept of criminal acts, criminal liability, and punishment.

1. Introduction

The rapid advancement and development of technology is a hallmark of the *Society 5.0* era, where technology and humans interact massively, integrating the virtual world and the physical world to create a better quality of life. In the development of digital technology, technologies such as Artificial Intelligence (AI), *Internet of Things* (IoT), *big data*, and fifth-generation (5G) and sixth-generation (6G) mobile networks have emerged.¹

Criminal law today is and for the next decades will faced with social changes and massive new technological developments in various fields. The social change in question is the way humans carry out their activities, from what was previously done conventionally, is now

¹ Maudy Andreana Lestari, Ahmad M Ramli, and Tasya Safiranita Ramli, "Telaah Yuridis Penyelenggaraan Teknologi 5G Di Indonesia: Langkah Transformasi Menuju Era Society 5.0," *CITIZENS: Jurnal Ilmiah Multidisiplin Indonesia* 2, no. 1 (2022): 129–37, 132, <https://doi.org/10.53866/jimi.v2i1.49>.

done digitally. In fact, social change continues to occur along with technological developments in various fields. One of the new technologies is artificial intelligence (AI). The presence of AI has started since the 1950s,² and since then, criminal law has not been able to adapt to its presence, even today. Is this situation due to the post-factum nature of criminal law?

This situation presents a dilemma in criminal law reform. The existence of AI is faced with a material criminal law regime that today is unable, perhaps even unwilling, to recognize and acknowledge the existence of AI. This is reasonable because AI is faced with three basic problems of material criminal law, namely criminal acts, criminal responsibility, and punishment. Even to get there, AI must first deal with the issue of its status, whether it can be made a legal subject or not.

Through this article, I would like to analyze the importance of the human role in realizing the existence of AI today. I believe that without it, criminal law will always lag behind social change and technological development, in line with the nature of law in general, which is always lagging. The role of humans is important because they are the ones who shape criminal law. The (criminal) law is presented for humans. This reminds me of an adage from Satjipto Rahardjo, that *"hukum ada untuk manusia, bukan manusia untuk hukum"*.

The presence of AI in social change and technological development also presents challenges in efforts to maintain peace, order, and justice for humans. These challenges can be in the form of new modus operandi of criminal acts. Just imagine if there is a criminal offense involving AI (with human victims) and the criminal law does not have a regulation or has not previously defined it. Isn't that tantamount to betraying the criminal law itself, which is present to protect the interests of the State, society, and individuals?

2. Discussion

2.1. Human and AI relationship

Today, Humans and AI coexist and need each other. Humans play a role in the development of AI. Conversely, AI exists and is developed to help humans with their tasks. This statement opens a discussion on the existence of humans and AI in the era of Industrial Revolution and Social Revolution. The existence of AI cannot be denied. AI is a part or result of the Fourth Industrial Revolution process. Klaus Schwab, founder of the World Economic Forum, defines this phase as an era characterized by a technological revolution that blurs the boundaries between the physical, digital and biological spheres.³ Before

² Cheny Berlian, Helmi Helmi, and Hafrida Hafrida, "The Urgency of Artificial Intelligence Criminal Responsibility as Cybercriminals," *International Journal of Scientific Multidisciplinary Research (IJSMR)* 2, no. 4 (2024): 297–310, <https://doi.org/10.55927/ijsmr.v2i4.8762>.

³ Klaus Schwab, "The Fourth Industrial Revolution: What It Means, How to Respond," World Economic Forum, 2016, <https://www.weforum.org/stories/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>; Philip Ross and Kasia Maynard, "Towards a 4th Industrial Revolution," *Intelligent Buildings International* 13, no. 3 (July 3, 2021): 159–61, 159, <https://doi.org/10.1080/17508975.2021.1873625>.

arriving at the Industrial Revolution 4.0, humans have been in contact with technology through the three previous phases of the industrial revolution.⁴

On the other hand, AI is reframing the relationship between humans and technology, and between humans and society. This is a social revolution characterized by Society 5.0. It combines cyberspace and physical space to balance economic progress with social problem solving through the provision of goods and services that elaborately address latent needs regardless of location, age, gender, or language. The vision of Society 5.0 is to reframe two types of relationships: the relationship between technology and society and the technology-mediated relationship between individuals and society.⁵

Those two revolutions above give us a glimpse of how science is focusing on AI and its relationship with humans. Not just on humans and society, and human-to-human relationships. This is a breakthrough in the way modern science is viewed. In the past, with science, humans managed to create AI; now, with science too, humans are trying to recognize and even interact with AI. So, does the interaction between humans and AI need to be regulated by law? Because the law exists to regulate relationships between individuals (humans and humans), between states (including between the state and legal entities/organizations/corporations), between individuals and the state (society), between individuals and legal entities/corporations, and among legal entities/corporations.⁶

The question and description above would invite us to reflect on the existence of AI and its relationship with humans using René Descartes' view of humans. We know that Descartes placed humans as one of the central issues in the history of his thought. Through Descartes, humans and their knowledge (reason/ratio) are used to find the truth.

3.2. Doubting [thinking] about the existence

René Descartes questioned the existence of man (himself) by doubting the existence of something, including man himself. According to him, when humans doubt something, it means that humans are thinking. For Descartes, the activity of doubting (thinking) aims to

⁴ Amr Adel, "Future of Industry 5.0 in Society: Human-Centric Solutions, Challenges and Prospective Research Areas," *Journal of Cloud Computing* 11, no. 1 (September 8, 2022): 40, <https://doi.org/10.1186/s13677-022-00314-5>. The First Industrial Revolution utilized water and steam power to mechanize production and infrastructure. The Second harnessed electricity to enable mass production. The Third marked the advent of personal computers, the internet, and the automation of production and operations beginning in the 1960s—later referred to as the "Digital Revolution." Currently, the Fourth Industrial Revolution builds upon its predecessors by leveraging the widespread availability of digital technologies developed during the Digital Revolution. It bridges the divide between the digital, physical, and biological domains, significantly accelerating the pace of technological change.

⁵ Atsushi Deguchi et al., "What Is Society 5.0?," in *Society 5.0* (Singapore: Springer Singapore, 2020), 1–23, 1, https://doi.org/10.1007/978-981-15-2989-4_1.

⁶ In the beginning, the law existed only to regulate the relationship between humans and between humans and the state (society). Then, with the rapid development in the economic field, the law recognized the existence of organizations and the like, which were regulated as legal subjects in the form of legal entities. Since then, law has recognized two concepts of legal subjects, individuals (human beings—*natuurlijk persons*) and groups of individuals/legal entities (corporations—*rechtspersons*).

get the ultimate truth. So that truth is obtained through knowledge (reason/ratio), not sensory experience. Therefore, humans must use reason (thinking) to find the truth. Descartes used the concept of thinking in a very broad sense. Something that thinks, according to him, is something that doubts, understands, affirms, rejects, wills, imagines, and feels, because the feeling when it appears in a dream is a form of thinking. Since thinking is the essence of the mind, the mind must always be thinking even when it is in deep sleep.⁷

Based on the above, I borrow Descartes' method of philosophizing to question the existence of humans and AI. I argue that by doubting the existence of humans and AI, we will realize the existence of humans and the existence of AI. By realizing the existence of both, we can realize the relationship between the two. Then what kind of relationship arises between both of them? However, before answering this question, we should first discuss the issue of the existence of humans and the existence of AI. Human existence is as a living being. Humans, in Descartes' view, are formed from two elements, namely the spirit and the body. The spirit is a single substance, which is immaterial and which cannot die. The soul has thought as its nature. What includes thought is everything that happens in humans with their knowledge, namely, all acts of sensory recognition, imagination, reason, and will. The essential nature of thought is consciousness.⁸ The body has as its fundamental property: vastness. As with all material things, the body is the object of physical science. Its actions are caused by its mechanical causes. There is an unbridgeable opposition between body and spirit. The apparent unity is only outward, as each embodies an independent thing. The essence of man is in his spirit. The spirit uses the body for certain actions. Although there is no meeting point between thought and expansiveness, the spirit and body influence each other as well.

AI's existence is as a living entity. A moving and tangible entity. Humans exist and live. Likewise, AI exists and lives. The issue that arises then is how the legal framework (in a general sense) and criminal law (in a specific sense) define the existence and life of AI. Of course, defining the existence and life of AI requires knowledge. Through the knowledge possessed, reason/ratio will guide humans to find answers—the ultimate truth. Since its birth in the 1950s,⁹ AI¹⁰ has been viewed in various states as a field of science, a tool, and an entity. In the view as a field of science, AI is a field devoted to building artifacts capable of displaying, in a controlled and well-understood environment, and over a sustained period, behaviors that we consider intelligent, or more generally, behaviors that we consider to be at the core of what is referred to as the mind.

⁷ Ngismatul Choiriyah, "Rasionalisme Rene Descartes," *Anterior Journal* 13, no. 2 (2014): 237–43, <https://media.neliti.com/media/publications/364622-none-8d511f30.pdf>.

⁸ Luh Putu Cita Ardiyani, Kadek Suciani, and Krisna Sukma Yogiswari, "Tubuh Dan Jiwa Manusia Dalam Filsafat Rene Descartes," *Vidya Darsan: Jurnal Mahasiswa Filsafat Hindu* 2, no. 2 (2021): 136–242, 139.

⁹ John McCarthy, "The Philosophy of AI and the AI of Philosophy," John McCarthy (Stanford, California, June 2006), <http://jmc.stanford.edu/articles/aiphil2.html>.

¹⁰ The term AI (Artificial Intelligence) was first introduced at the Dartmouth Workshop in 1956. This moment became the starting point for the birth of a new field called artificial intelligence.

As John McCarthy cites Eugene Charniak and Drew McDermott, AI is a field devoted to building people.¹¹ Charniak and McDermott argue that the ultimate goal of AI, which we are very far from achieving, is to build a person or, more humbly, an animal.¹² This view is supported by John Haugeland, who states that the fundamental goal [of AI research] is not merely to mimic intelligence or produce some clever fake. Not at all. AI wants only the genuine article: machines with minds, in the full and literal sense. This is not science fiction, but real science, based on a theoretical conception as deep as it is daring: namely, we are, at root, computers ourselves.¹³

It can be seen that the fundamental goal [of AI research] is not just to mimic intelligence or produce something smart. Not at all. AI only wants genuine capabilities: machines that have minds, in the true sense of the word. This is not science fiction, but real science, based on theoretical conceptions as deep and bold as: that is, we are essentially computers ourselves. In the tool view, AI is seen as a tool that functions to perform specific tasks, such as data processing, pattern recognition, and machine learning. AI performs its functions by relying on hardware and software designed by humans. As of today, AI performs functions and relies on human control in terms of developing, using, and regulating AI. In the field of AI studies, this situation is part of what is known as Artificial Narrow Intelligence (ANI).¹⁴ The ability of AI to perform a complex task based on commands (programs) created by humans.

In the entity view, AI can learn from data and experience. The data learned comes from Big Data. With machine learning (ML) and deep learning (DL) capabilities, AI will learn from the data and the learning process, and AI forms experiences.¹⁵ This picture allows AI to adjust to new situations—new data—inputted into Big Data. The new data is learned and assessed by AI. The assessment then produces an output in the form of a decision to do or not do something. AI's learning and adaptation capabilities enable it to interact with humans and the environment.

Returning to Descartes' view, it can be realized that the existence of humans as humans is because humans are formed from the body and spirit. The essence of man is in his soul. The soul uses the body for certain actions. Meanwhile, the existence of AI is realized as a

¹¹ McCarthy, "The Philosophy of AI and the AI of Philosophy"; Eugene Charniak and Drew McDermott, *Introduction to Artificial Intelligence* (Reading, Mass.: Addison-Wesley, 1987).

¹² Charniak and McDermott, *Introduction to Artificial Intelligence*.

¹³ John Haugeland, *Artificial Intelligence: The Very Idea* (Cambridge: MIT Press, 1989).

¹⁴ Paolo Bory, "Deep New: The Shifting Narratives of Artificial Intelligence from Deep Blue to AlphaGo," *Convergence: The International Journal of Research into New Media Technologies* 25, no. 4 (August 12, 2019): 627–42, <https://doi.org/10.1177/1354856519829679>; Ragnar Fjelland, "Why General Artificial Intelligence Will Not Be Realized," *Humanities and Social Sciences Communications* 7, no. 1 (June 17, 2020): 10, <https://doi.org/10.1057/s41599-020-0494-4>; V. Sowri Rabu and Krishna Banana, "A Study on Narrow Artificial Intelligence - An Overview," *International Journal of Engineering Science and Advanced Technology* 24, no. 4 (2024): 210–19, https://ijesat.com/ijesat/files/V24I0428_1714383466.pdf.

¹⁵ Anitha S. Pillai and Roberto Tedesco, "Introduction to Machine Learning, Deep Learning, and Natural Language Processing," in *Machine Learning and Deep Learning in Natural Language Processing*, ed. Anitha S. Pillai and Roberto Tedesco (Boca Raton: CRC Press, 2023), 3–14, <https://doi.org/10.1201/9781003296126-2>.

tool and or entity. AI performs certain tasks, such as data processing, pattern recognition, and machine learning. AI performs its functions by relying on hardware and software designed by humans. AI performs functions and relies on human control. It is also important to realize that as an entity, it can be said that AI is similar to humans. AI does learning activities and has experience (from learning about the data contained in Big Data). Like humans, AI can also, or at least learn new things (such as situations) to adjust and adapt to change. In addition, AI also goes through the same decision-making process as humans. It receives information, manages it, and produces outputs in the form of autonomous decisions and/or actions.

Realizing the existence of AI as a tool and entity shows that AI only has a body, without a spirit. Unlike humans, who have both. Thus, it can be concluded that life in the sense of being human is having a body and a spirit. Life in the sense of being an AI (tool or entity) is only having a body, without a spirit. However, the explanation that has been given gives rise to new doubts. Doubts about whether AI will one day have a spirit? Or, if it is possible, can spirit be defined in another sense, different from how it is commonly defined today? Of course, these doubts are questions that need to be answered by thinking humans.

3.3. Three-dilemmatic issues

Realizing the existence of humans and the existence of AI in social change and technological development is a dilemma for criminal law reform. As is known, criminal law reform is aimed at updating the concept of three basic problems of criminal law, namely criminal offense, criminal liability, and punishment (hereinafter referred to as the Three Issues). For me, this is the three-dilemmatic issues faced, in particular, with the massive development of AI—among others: AI and criminal offenses, AI and criminal liability, and AI and punishment.

First Issue, AI and criminal offenses (hereinafter referred to as “First Issue”). The First Issue relates to the principle of legality in criminal law and the definition of criminal offense. The principle of legality, *Nullum delictum, nulla poena sine praevia lege poenali*, which means no act is punishable except under the criminal provisions in the laws and regulations in force before the act was committed.¹⁶ This principle was first formulated by a German criminal law expert, Paul Johan Anslem von Feuerbach (1775-1833), in his book entitled *Lehrbuch des Penlichen Recht* in 1801.¹⁷ Bambang Poernomo argues that

¹⁶ Beni Puspito and Ali Masyhar, “Dynamics of Legality Principles in Indonesian National Criminal Law Reform,” *Journal of Law and Legal Reform* 4, no. 1 (January 22, 2023): 129–148, <https://doi.org/10.15294/jllr.v4i1.64078>; Jan Remmelink, *Hukum Pidana: Komentar Atas Pasal-Pasal Terpenting dari Kitab Undang-Undang Hukum Pidana Belanda dan Padanannya dalam Kitab Undang-Undang Hukum Pidana Indonesia* (Jakarta: Gramedia Pustaka Utama, 2003).

¹⁷ Bambang Poernomo, “Manfaat Telaah Ilmu Hukum Pidana Dalam Membangun Model Penegakan Hukum Pidana Di Indonesia” (Yogyakarta: Univesitas Gajah Mada, 1989).

Feuerbach formulated the principle of legality with a note; the principle of legality aims to protect individual interests, which is characteristic of the purpose of criminal law.¹⁸

According to George P. Fletcher, the principle of legality has two meanings, a positive and a negative one. In a positive sense, the principle of legality serves to protect individuals from arbitrary actions of the State.¹⁹ In a negative sense, the principle of legality serves to protect society by punishing criminals who are declared [verdict] guilty by the State. Another definition was conveyed by Moeljatno, that no act is prohibited and threatened with punishment if it is not determined in advance in the legislation. Furthermore, he formulated the principle of legality in three senses:²⁰ 1) An act cannot be prohibited and punishable if the act has not been determined in advance in a statutory regulation; 2) To determine the existence of a criminal offences, analogies cannot be used; dan 3) The provisions of criminal law do not apply retroactively.

Keep in mind that the principle of legality is related to time. The interpretation of the principle of legality of criminal law is the application of criminal law or a criminal law regulation related to time. If an act that fulfills the formulation of a criminal offense committed before the enactment of the relevant provisions, then not only the act cannot be prosecuted, but also cannot be blamed for the act at all. In accordance with its formulation, the principle of legality requires the prior regulation of an act in the criminal law regulations, before blaming the person who committed it.²¹ Traditionally, the principle of legality demands clarity, certainty, and the prohibition of retroactivity.²² However, in modern practice, this principle faces challenges to adapt to the dynamics of a rapidly developing society, such as transnational crimes, technology-based crimes, and the application of international law. This challenge has led to a discourse on the flexibility of the application of the principle of legality without compromising its basic principles, namely the protection of human rights and legal certainty.²³

Today, the above definition is faced with the existence of AI as a tool and/or entity. So, if a criminal offense occurs and AI plays a role in it, can it still be called a criminal offense? Of course, the answer is yes, a criminal offense, as long as the act committed has been

¹⁸ Remmelink, *Hukum Pidana: Komentar Atas Pasal-Pasal Terpenting Dari Kitab Undang-Undang Hukum Pidana Belanda Dan Padanannya Dalam Kitab Undang-Undang Hukum Pidana Indonesia*. Aside from individuals, criminal law also aims to protect the interests of the State and society.

¹⁹ George P. Fletcher, *Basic Concepts of Criminal Law* (New York: Oxford University Press, 1998).

²⁰ Moeljatno, *Asas-Asas Hukum Pidana*, Revisi (Jakarta: Rineka Cipta, 2008).

²¹ Danel Aditia Situngkir, "Asas Legalitas Dalam Hukum Pidana Nasional Dan Hukum Pidana Internasional," *Soumatara Law Review* 1, no. 1 (2018): 22-42, <https://doi.org/10.22216/soumlaw.v1i1.3398>; Vincentius Patria Setyawan, "Asas Legalitas Dalam Perspektif Filsafat Hukum," *Justitia et Pax* 37, no. 1 (2021): 127-46, <https://jurnal.hukumonline.com/a/62360de517eef054e47933bb/asas-legalitas-dalam-perspektif-filsafat-hukum>.

²² F.A. Fitri et al., "Tinjauan Teoritis Tentang Asas Legalitas Dalam Hukum Pidana Indonesia," *JIMMI: Jurnal Ilmiah Mahasiswa Multidisiplin* 1, no. 2 (2024): 202-9, <https://doi.org/10.71153/jimmi.v1i2.134>.

²³ Dedi Iskandar et al., "Perkembangan Teori Dan Penerapan Asas Legalitas Dalam Hukum Pidana Indonesia," *JIMMI: Jurnal Ilmiah Mahasiswa Multidisiplin* 1, no. 3 (2024): 293-305, <https://doi.org/doi.org/10.71153/jimmi.v1i3.147>.

criminalized.²⁴ The focus of Issue One is the criminal offense²⁵ (an unlawful act prohibited by criminal law to be done or not done), not the perpetrator. However, by act, we mean an act committed by a human being, not another entity, such as an AI.

Herein lies the dilemma in Issue One. If the existence of AI is seen (realized) as a tool, then the criminal offense that occurs remains a criminal offense, and AI is only seen as a tool used/utilized to commit a criminal offense. However, if the existence of AI is seen (realized) as an entity, shouldn't criminal offenses also be interpreted as criminal offenses committed by AI? This question is raised to challenge the notion of criminal offense in its current understanding, which is only interpreted as an act committed by humans.

There are various definitions of criminal offense, as quoted from Amir Ilyas.²⁶ Some of them are, as stated by Moeljatno and Pompe. According to Moeljatno, who uses the term "*perbuatan pidana*", it is stated that a criminal offense is an act that violates which is prohibited by a rule of law, which prohibition is accompanied by a threat (sanction) in the form of a certain punishment for anyone who violates the prohibition. Pompe defines criminal offenses in two ways. First, the theoretical definition–violation of norms (rules and legal system), which is held due to the fault of the violator, and must be given a penalty to be able to maintain the legal system and save public welfare. Second, a statutory definition–an event that is determined by law to contain action (*handeling*) and neglect (*nalaten*); inaction; passive action, usually carried out in several circumstances that are part of an event.

The definition of a criminal offense is composed of elements that make it a criminal offense. E.Y. Kanter and S.R. Sianturi formulated five elements that form the definition of a criminal offense, among others:²⁷ 1) subject; 2) fault; 3) unlawful nature of an act; 4) an act that is prohibited or required by law and against which the offender is punished; and 5) time, place, and circumstances (other objective elements). These five elements can be reformulated into only two elements, subjective and objective. The subjective element consists of "subject" and "fault"–relating to the *mens rea* (evil intent) of the perpetrator. Meanwhile, the objective elements consist of "the unlawful nature of the act", "an act

²⁴ BPHN, "Draft Naskah Akademik Rancangan Undang-Undang Tentang Kitab Undang-Undang Hukum Pidana (KUHP)" 2015 (Jakarta, 2015). Muladi reminded that criminalization and decriminalization must be guided by the following 8 things: 1) criminalization must not effectively cause "overcriminalization" which is categorized as "the misuse of criminal sanction"; 2) criminalization must not be *ad hoc*; 3) criminalization must contain elements of victims, both actual and potential; 4) criminalization must consider cost and outcome analysis (cost benefit principle); 5) criminalization must obtain public support; 6) criminalization must result in "enforceable" regulations – law enforcement can be carried out; 7) criminalization must contain elements of subsocialitiet (resulting in harm to society even if it is very small); and 8) criminalization must heed the warning that any criminal regulation restricts the freedom of the people and gives law enforcement officials the possibility to curb that freedom.

²⁵ Criminal acts as stipulated in Article 12 of Law Number 1 of 2023 concerning the Criminal Code.

²⁶ Amir Ilyas, *Asas-Asas Hukum Pidana: Memahami Tindak Pidana Dan Pertanggungjawaban Pidana Sebagai Syarat Pemidanaan*, ed. Andi Maulana Mustamin (Yogyakarta: Mahakarya Rangkang Offset Yogyakarta, 2012).

²⁷ E.Y. Kanter and S.R. Sianturi, *Azaz-Azaz Hukum Pidana Di Indonesia Dan Penerapannya* (Jakarta: Stora Grafika, 2002).

prohibited or required by law and against which punishment is imposed”, and “time, place, and circumstances”–relating to the *actus reus* (malicious act) of the perpetrator.

This definition of criminal offense is understandable because conceptually, the elements of *actus reus* (evil act) and *mens rea* (evil intent) play a role in the criminalization of the perpetrator of a criminal offense, with a note: if there is no justification. These two elements are at the core of the discussion of the two main approaches in criminal law, namely dualism and monism. Dualism strictly separates *actus reus* and *mens rea*. Monism, on the other hand, does not. Monism emphasizes that the main focus should be on the act itself. When the prohibited act is proven, then the mental attitude is assumed to have been fulfilled so that the perpetrator can be reproached and held responsible.²⁸

The definition of criminal offense and the issue of legality become a dilemma for criminal law reform when it is realized that the existence of AI can be seen as an entity that in its development, will or potentially have the same capabilities as humans. Therefore, in such a situation, the existing definition of criminal offense should be questionable for the future. Therefore, the question arises: Can the definition of criminal offense also be formulated in terms of acts committed by AI?

The illustration presented in the First Issue brings us to a second, more complex dilemma–namely, the issue of artificial intelligence (AI) and criminal liability (hereinafter referred to as the “Second Issue”). In essence, the Second Issue has been extensively examined in scholarly discourse concerning the attribution of criminal responsibility to AI. As a matter of principle, a perpetrator may be held criminally liable for a criminal offense insofar as the offense satisfies two fundamental elements: *actus reus* (evil act) and *mens rea* (evil intent).

Referring back to the distinction between monist and dualist perspectives, a critical question arises as to which of these theoretical frameworks may be most appropriate for formulating a concept of criminal liability in the context of artificial intelligence (AI). It is important to recall that a fundamental component inherent in the concept of criminal liability is culpability. *Geen straf zonder schuld*–no punishment without fault. This principle positions culpability as a foundational element in determining whether a perpetrator of a criminal act may justly be held blameworthy under criminal law.

Herein lies the core dilemma of the Second Issue. I remain skeptical as to whether artificial intelligence (AI) is capable of possessing culpability. To date, the element of culpability has been attributed exclusively to individual perpetrators–natural persons. It was only later, with the recognition of corporations as legal subjects (*rechtspersoon*), that culpability was extended to corporate entities. It is this attribution of culpability that enables a legal

²⁸ Zul Khaidir Kadir, “Dari Dualisme Ke Monisme: Transformasi Konsep Mens Rea Dalam Kodifikasi KUHP Di Negara-Negara Poskolonial,” *JULIA: Jurnal Litigasi Amsir* September-, no. Spesial Isu (2024), <https://journalstih.amsir.ac.id/index.php/julia/article/view/605>.

subject to bear criminal liability.²⁹ This raises a critical question: can AI likewise possess culpability in a manner that would render it subject to criminal liability?³⁰

Returning to the issue of culpability, it is generally classified into two forms: intent (*dolus*) and negligence (*culpa*). Concerning intent, two principal theories have been developed—namely, the theory of will and the theory of knowledge. According to the theory of will, intent is defined as the deliberate will directed toward the realization of an objective as prohibited by law, emphasizing the prohibited goal. In contrast, the theory of knowledge holds that intent consists of the will to act while being aware of the elements that constitute a criminal offense under the law, emphasizing the prohibited act.

Can it be provisionally concluded that AI is capable of recognizing the commission of a criminal offense? While such a conclusion may be subject to doubt, it may nonetheless serve as a basis for re-examining the issue of whether AI can possess culpability in the context of a criminal act. Accordingly, the absence of culpability in AI would suggest that it cannot be held criminally liable. However, this conclusion remains contestable. We are faced with the broader question of whether the concept of mens rea, as it exists today in the digital era, has become inherently problematic. On the other hand, an offender who possesses culpability is presumed to have the capacity to be held accountable. E.Y. Kanter and S.R. Sianturi argue that the capacity for criminal responsibility is grounded in the condition and capability of the “spirit,” rather than in one’s mere capacity to “think.”³¹

Since AI is not a legal subject and does not possess a soul, the imposition of criminal liability—as previously discussed, becomes inherently problematic. Nevertheless, it is useful to consider the concept of criminal liability for AI as formulated by Gabriel Hallevy. According to Hallevy, criminal offenses involving AI may arise under three possible scenarios. First, an offense may occur due to the deliberate intent of a legal subject to utilize AI for unlawful purposes. Second, an offense may result from the intentional use of AI by a legal subject to avoid fulfilling a legal obligation. Third, an offense may arise from the negligence of a legal subject in the use of AI. Hallevy proposes three models of liability

²⁹ Fifink Praiseda Alviolita, “Pertanggungjawaban Pidana Korporasi Oleh Pengurus Korporasi Dikaitka Dengan Asas Geen Straf Zonder Schuld,” *Refleksi Hukum: Jurnal Ilmu Hukum* 3, no. 1 (May 5, 2019): 1–16, <https://doi.org/10.24246/jrh.2018.v3.i1.p1-16>. In terms of corporate liability, it is known as the principle of strict liability which is a doctrine of corporate criminal liability adopted from the doctrine of civil law, which is the doctrine applied in unlawful acts. In the event that the corporation violates or does not fulfill certain obligations/conditions/situations by this corporation is known as “companies offence”, “situational offence”, or “strict liability offences”

³⁰ Criminal liability is only imposed on perpetrators who have mistakes in criminal acts. As for errors, they are only inherent in legal subjects, and AI is not legal subjects. Of course, the first thing that must be passed is that AI is recognized as a legal subject. However, to get there, it may still take a few more decades as criminal law realizes the massive development of AI (in legal traffic and technology). It is the same when criminal law is aware of the development of corporations in legal and economic traffic).

³¹ Kanter and Sianturi, *Azaz-Azaz Hukum Pidana Di Indonesia Dan Penerapannya*. E.Y. Kanter and S.R. Sianturi formulated that a person can be charged with criminal liability if the act is unlawful (there is no elimination of unlawful nature and/or the existence of a justifiable reason). As for Kanter and Sianturi, they formulated two elements of responsible ability: the state of the soul (the perpetrator) and the ability of the soul (the perpetrator). These two elements are souls, and AI has no souls. This brings us back to Descartes’ dualistic view of human beings being composed of two elements, the body and the spirit.

for offenses committed by AI: the Perpetration-via-Another Model (PVM), the Natural-Probable Consequence Liability Model (NPCLM), and the Direct Liability Model (DLM).³²

The first model, Perpetration-via-Another Model (PVM), views AI as an agent lacking the capacity to commit a criminal offense independently. In this model, AI functions merely as an instrument or tool, while the actual perpetrator is the individual involved in the development or use of the AI system that results in a criminal act. Criminal liability is therefore attributed to the human actor behind the AI. This could be a programmer who intentionally designs the AI to commit offenses under specific circumstances, or an owner/user who deliberately employs the AI for criminal purposes. This accountability approach is grounded in strict liability, meaning that criminal responsibility is imposed without the need to assess intent or negligence. The rationale is that the deliberate creation and use of AI for unlawful objectives constitutes a significant danger to society, justifying the imposition of liability regardless of the AI's autonomous status.

The second model, Natural-Probable Consequence Liability Model (NPCLM), focuses on criminal offenses that arise due to faults or malfunctions within the AI system. For example, consider the case in 1981 when a robot killed a worker at a motorcycle factory in Japan.³³ The robot identified the worker as an obstacle to its task and, based on its embedded AI system, calculated the necessary action to remove the worker by pushing him into an operating machine. In such cases, criminal liability may be directly imposed on all responsible parties (owners, designers, or supervisors) if they were aware of the potential risks involved. Liability may also arise if these parties failed to recognize weaknesses in the AI system they employed. Furthermore, proportional liability can be applied when the causal links and consequences vary among different manufacturers or operators involved in the offense.

The third model, Direct Liability Model (DLM), treats AI similarly to a human being in terms of criminal liability. In this model, AI is not considered an innocent agent but is attributed with a certain level of consciousness or mental state at the time a criminal act occurs. In other words, AI is deemed capable of bearing responsibility. At first glance, the DLM appears difficult to conceptualize, as assigning mental attributes to AI presents a significant challenge in many cases. For example, if a self-driving car exceeds the legally prescribed speed limit, the AI technology could be deemed to have violated its duty of care. The validity of this approach is widely debated. Questions arise, such as how to address

³² Asri Gresmelian Eurike Hailtik and Wiwik Afifah, "Criminal Responsibility of Artificial Intelligence Committing Deepfake Crimes in Indonesia," *Asian Journal of Social and Humanities* 2, no. 4 (2024): 776–95, <https://doi.org/10.59888/ajosh.v2i4.222>; Gabriel Hallevy, "Dangerous Robots–Artificial Intelligence vs. Human Intelligence," SSRN, 2018, <https://doi.org/10.2139/ssrn.3121905>.

³³ Gabriel Hallevy, *When Robots Kill: Artificial Intelligence Under Criminal Law* (Boston: Northeastern University Press, 2013).

liability violations resulting from software errors or AI actions performed in self-defense, which might be analogous to doctrines of justification or excuse under criminal law.³⁴

The three model scenarios formulated by Hallevy serve as a bridge to the third dilemma, namely AI and criminal punishment (hereinafter referred to as the Third Issue). Referring to the scenarios above, it can be argued that AI may be subject to criminal sanctions. The question then arises: What type of punishment would be applicable? Should existing forms of punishment, which are typically imposed on human individuals or corporate entities, be applied to AI? This constitutes the core dilemma of the Third Issue. Punishment is a condemnation (sanction) imposed by the state (representing society) on perpetrators of criminal offenses, whether individuals or corporations, who have been found guilty of committing crimes. The types and forms of punishment imposed on offenders are guided by the objectives of criminal sanctions. Accordingly, the question arises: for what purpose would the state punish AI that commits (or is involved in) a criminal offense? The formulation of these punitive objectives was based on the premise that punishment, in essence, is merely an instrument to achieve certain goals.

Referring to the third model, the Direct Liability Model (DLM)—where AI is treated the same as humans in criminal liability—it can be argued that AI possesses culpability, making it subject to punishment if it commits a criminal offense. In this context, punishment serves as a tool to penalize AI. However, treating AI the same as humans raises the question of whether AI should also share the same purposes of punishment. Herbert L. Packer states that the retributive perspective is based on the idea that it is just for wrongdoers to be punished: Since humans are responsible for their actions, they must receive deserved retribution.³⁵ This view is grounded in the objective of providing just deserts to offenders for their criminal conduct.

If one were to reinterpret and reformulate Packer’s assertion through the lens of criminal law theory, it might read: “The retributive theory is premised on the notion that it is just to punish a malevolent AI: because the AI bears responsibility for its actions, it ought to receive a proportionate punishment.” Within this formulation, the prospect of subjecting AI to criminal liability becomes theoretically conceivable. Punishment is imposed not merely for the existence of AI, which consists of a physical embodiment devoid of spirit or consciousness, but for its capacity to engage in autonomous decision-making. In contemporary contexts, AI systems—although still reliant on human programming—demonstrate the ability to make independent choices, including the decision to act or refrain from acting. If criminal responsibility is predicated on conduct rather than cognition, the focus shifts from intent to behavior. Thus, a critical question arises: would

³⁴ Indeed, granting AI status as a legal subject means involving all parties involved in the planning, manufacturing process, implementation, or ownership of the technology to be held criminally liable regardless of the proportion of the part given.

³⁵ Herbert L. Packer, *The Limits of Criminal Sanctions* (Stanford, California: Stanford University Press, 1978).

subjecting an AI to punishment constitute the fair and just retribution contemplated by penal theory?

However, with the reform of criminal law, the objective of punishment has evolved beyond mere retribution. Of greater importance now is the aim of deterrence. In this context, punishment is imposed not only to protect the interests of the state, society (particularly victims), and individuals, but also to safeguard the offender's interests. This raises a critical question in the context of artificial intelligence: Does AI possess any interests worthy of legal protection? In discussing the aims of punishment, Rupert Cross identifies three fundamental justifications: prevention, deterrence, and reform.³⁶

First, prevention; refers to rendering the offender incapable of committing further crimes (incapacitation). This incapacitation may be either short-term or long-term. Historically, the most effective forms of punishment in this regard have been the death penalty and life imprisonment. While such measures are considered justifiable, they are typically reserved for the most serious offenses. One might imagine a scenario in which an AI system commits a grave crime, and punishment is imposed for the purpose of prevention. In such a case, the AI could be "executed"—its entire system permanently shut down.

Second, deterrence; typically categorized into individual deterrence, general deterrence, and long-term deterrence, is based on the rationale that the experience of punishment, the threat of sanctions, or exemplary punishment can suppress criminal behavior. Individual deterrence is grounded in the belief that the offender must undergo a punitive experience to discourage future offenses. In this context, punishment is imposed to make the AI "reluctant" to reoffend. This raises a crucial question: can AI become a recidivist? Next is general deterrence, which is based on the expectation that the punishment of an individual offender will serve as an example, thereby discouraging others from committing similar crimes. Yet, one must ask: can AI experience fear? Following this is long-term deterrence, which refers to the hope that the immediate punitive impact will influence the offender's future conduct. However, Cross argues that such a scope is insufficient and must be complemented by denunciatory and educative elements, grounded in the effort to establish behavioral standards for human society. I am of the view that these denunciatory and educative elements can be operationalized through programmers who reprogram AI systems to align with human-friendly behavior.

Finally, the third objective is reform. Cross distinguishes between "the idea that reform can come through the punishment itself", which he considers outdated, and "the idea of reform as a concomitant of punishment", which he views as modern. Cross critiques the earlier notion that punishment alone can bring about reform, arguing that such a view is no longer suitable in the contemporary era and should be replaced with the idea of reform as an accompanying element of punishment. In this light, punishment imposed on AI is no

³⁶ Rupert Cross, *The English Sentencing System* (London: Butterworths, 1975).

longer merely retributive, but may also serve as a means of reform—that is, to improve or reinstall the AI system.

The uncertainty surrounding the Third Issue is particularly complex. If AI can be held criminally liable under the Direct Liability Model (DLM)—where AI is treated similarly to human beings in terms of criminal responsibility—then it becomes conceivable for AI to be subjected to punishment. The types of punishment that could be imposed on AI vary, including the death penalty in the form of a system takedown, particularly in cases involving serious crimes. However, the death penalty could potentially be substituted with other forms of punishment, such as reforming, upgrading, or reinstalling the AI system.

3. Conclusion

The massive development of technology in today's digital era makes the relationship between humans and AI unavoidable. The relationship must be based on the awareness of the existence of humans and AI that both exist and live. However, when using René Descartes' moral view of the idea of world (elemental) dualism, it can be concluded that humans and AI exist and live in different existences. Humans exist and live as humans because they have two constituent elements, namely the body and the spirit. AI, on the other hand, exists and lives as a tool and or entity that only has a body, without a spirit. Realizing the existence of AI only raises doubts. can AI also have a spirit, or if possible, can the spirit be redefined?

Realizing the existence of AI, which only has a body, without a spirit, presents a dilemma in criminal law reform. The dilemma is present because today, AI continues to experience massive developments. Of course, this massive development makes me doubt the concepts of the three basic problems of criminal law, namely criminal acts, criminal responsibility, and punishment. These doubts need to be criticized by re-questioning the definitions of the three basic issues of criminal law. The definitions of the three should also mean that AI can commit a criminal offense, can have guilt, and ultimately can be punished. These definitions become a homework assignment for criminal law reform in the future. The point is the next studies should address the redefined on artificial intelligence and the three fundamental issues of criminal law itself based on criminal policy and criminology through ethics studies.

Daftar Pustaka

- Adel, Amr. "Future of Industry 5.0 in Society: Human-Centric Solutions, Challenges and Prospective Research Areas." *Journal of Cloud Computing* 11, no. 1 (September 8, 2022): 40. <https://doi.org/10.1186/s13677-022-00314-5>.
- Alviolita, Fifink Praiseda. "Pertanggungjawaban Pidana Korporasi Oleh Pengurus Korporasi Dikaitkan Dengan Asas Geen Straf Zonder Schuld." *Refleksi Hukum: Jurnal Ilmu Hukum* 3, no. 1 (May 5, 2019): 1–16. <https://doi.org/10.24246/jrh.2018.v3.i1.p1-16>.
- Ardiyani, Luh Putu Cita, Kadek Suciani, and Krisna Sukma Yogiswari. "Tubuh Dan Jiwa Manusia Dalam Filsafat Rene Descartes." *Vidya Darsan: Jurnal Mahasiswa Filsafat Hindu*

2, no. 2 (2021): 136–242, 139.

Berlian, Cheny, Helmi Helmi, and Hafrida Hafrida. "The Urgency of Artificial Intelligence Criminal Responsibility as Cybercriminals." *International Journal of Scientific Multidisciplinary Research (IJSMR)* 2, no. 4 (2024): 297–310. <https://doi.org/10.55927/ijsmr.v2i4.8762>.

Bory, Paolo. "Deep New: The Shifting Narratives of Artificial Intelligence from Deep Blue to AlphaGo." *Convergence: The International Journal of Research into New Media Technologies* 25, no. 4 (August 12, 2019): 627–42. <https://doi.org/10.1177/1354856519829679>.

BPHN. "Draft Naskah Akademik Rancangan Undang-Undang Tentang Kitab Undang-Undang Hukum Pidana (KUHP)." 2015. Jakarta, 2015.

Charniak, Eugene, and Drew McDermott. *Introduction to Artificial Intelligence*. Reading, Mass.: Addison-Wesley, 1987.

Choiriyah, Ngismatul. "Rasionalisme Rene Descartes." *Anterior Journal* 13, no. 2 (2014): 237–43. <https://media.neliti.com/media/publications/364622-none-8d511f30.pdf>.

Cross, Rupert. *The English Sentencing System*. London: Butterworths, 1975.

Deguchi, Atsushi, Chiaki Hirai, Hideyuki Matsuoka, Taku Nakano, Kohei Oshima, Mitsuharu Tai, and Shigeyuki Tani. "What Is Society 5.0?" In *Society 5.0*, 1–23, 1. Singapore: Springer Singapore, 2020. https://doi.org/10.1007/978-981-15-2989-4_1.

Fitri, F. A., N. Muftia, I. Trilia, A. H. Munthe, and R. Ramlan. "Tinjauan Teoritis Tentang Asas Legalitas Dalam Hukum Pidana Indonesia." *JIMMI: Jurnal Ilmiah Mahasiswa Multidisiplin* 1, no. 2 (2024): 202–9. <https://doi.org/10.71153/jimmi.v1i2.134>.

Fjelland, Ragnar. "Why General Artificial Intelligence Will Not Be Realized." *Humanities and Social Sciences Communications* 7, no. 1 (June 17, 2020): 10. <https://doi.org/10.1057/s41599-020-0494-4>.

Fletcher, George P. *Basic Concepts of Criminal Law*. New York: Oxford University Press, 1998.

Hailtik, Asri Gresmelian Eurike, and Wiwik Afifah. "Criminal Responsibility of Artificial Intelligence Committing Deepfake Crimes in Indonesia." *Asian Journal of Social and Humanities* 2, no. 4 (2024): 776–95. <https://doi.org/10.59888/ajosh.v2i4.222>.

Hallevy, Gabriel. "Dangerous Robots – Artificial Intelligence vs. Human Intelligence." SSRN, 2018. <https://doi.org/10.2139/ssrn.3121905>.

———. *When Robots Kill: Artificial Intelligence Under Criminal Law*. Boston: Northeastern University Press, 2013.

Haugeland, John. *Artificial Intelligence: The Very Idea*. Cambridge: MIT Press, 1989.

Ilyas, Amir. *Asas-Asas Hukum Pidana: Memahami Tindak Pidana Dan Pertanggungjawaban Pidana Sebagai Syarat Pemidanaan*. Edited by Andi Maulana Mustamin. Yogyakarta: Mahakarya Rangkang Offset Yogyakarta, 2012.

Iskandar, Dedi, W.N. Zulbaidah, Angga Almada, Iswandi Abdinur, Devi Yanda Putra, Cut Yessi Andriani, and Zulhazrul. "Perkembangan Teori Dan Penerapan Asas Legalitas

- Dalam Hukum Pidana Indonesia." *JIMMI: Jurnal Ilmiah Mahasiswa Multidisiplin* 1, no. 3 (2024): 293–305. <https://doi.org/doi.org/10.71153/jimmi.v1i3.147>.
- Kadir, Zul Khaidir. "Dari Dualisme Ke Monisme: Transformasi Konsep Mens Rea Dalam Kodifikasi KUHP Di Negara-Negara Poskolonial." *JULIA: Jurnal Litigasi Amsir* September-, no. Spesial Isu (2024). <https://journalstih.amsir.ac.id/index.php/julia/article/view/605>.
- Kanter, E.Y., and S.R. Sianturi. *Azaz-Azaz Hukum Pidana Di Indonesia Dan Penerapannya*. Jakarta: Stora Grafika, 2002.
- Lestari, Maudy Andreana, Ahmad M Ramli, and Tasya Safiranita Ramli. "Telaah Yuridis Penyelenggaraan Teknologi 5G Di Indonesia: Langkah Transformasi Menuju Era Society 5.0." *CITIZENS: Jurnal Ilmiah Multidisiplin Indonesia* 2, no. 1 (2022): 129–37. <https://doi.org/10.53866/jimi.v2i1.49>.
- McCarthy, John. "The Philosophy of AI and the AI of Philosophy." John McCarthy. Stanford, California, June 2006. <http://jmc.stanford.edu/articles/aiphil2.html>.
- Moeljatno. *Asas-Asas Hukum Pidana*. Revisi. Jakarta: Rineka Cipta, 2008.
- Packer, Herbert L. *The Limits of Criminal Sanctions*. Stanford, California: Stanford University Press, 1978.
- Pillai, Anitha S., and Roberto Tedesco. "Introduction to Machine Learning, Deep Learning, and Natural Language Processing." In *Machine Learning and Deep Learning in Natural Language Processing*, edited by Anitha S. Pillai and Roberto Tedesco, 3–14. Boca Raton: CRC Press, 2023. <https://doi.org/10.1201/9781003296126-2>.
- Poernomo, Bambang. "Manfaat Telaah Ilmu Hukum Pidana Dalam Membangun Model Penegakan Hukum Pidana Di Indonesia." Yogyakarta: Univesitas Gajah Mada, 1989.
- Puspito, Beni, and Ali Masyhar. "Dynamics of Legality Principles in Indonesian National Criminal Law Reform." *Journal of Law and Legal Reform* 4, no. 1 (January 22, 2023): 129–48. <https://doi.org/10.15294/jllr.v4i1.64078>.
- Rabu, V. Sowri, and Krishna Banana. "A Study on Narrow Artificial Intelligence - An Overview." *International Journal of Engineering Science and Advanced Technology* 24, no. 4 (2024): 210–19. https://ijesat.com/ijesat/files/V24I0428_1714383466.pdf.
- Rommelink, Jan. *Hukum Pidana: Komentor Atas Pasal-Pasal Terpenting Dari Kitab Undang-Undang Hukum Pidana Belanda Dan Padanannya Dalam Kitab Undang-Undang Hukum Pidana Indonesia*. Jakarta: Gramedia Pustaka Utama, 2003.
- Ross, Philip, and Kasia Maynard. "Towards a 4th Industrial Revolution." *Intelligent Buildings International* 13, no. 3 (July 3, 2021): 159–61, 159. <https://doi.org/10.1080/17508975.2021.1873625>.
- Schwab, Klaus. "The Fourth Industrial Revolution: What It Means, How to Respond." World Economic Forum, 2016. <https://www.weforum.org/stories/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>.
- Setyawan, Vincentius Patria. "Asas Legalitas Dalam Perspektif Filsafat Hukum." *Justitia et Pax* 37, no. 1 (2021): 127–46. <https://jurnal.hukumonline.com/a/62360de517eef054e47933bb/asas-legalitas->

dalam-perspektif-filsafat-hukum.

Situngkir, Danel Aditia. "Asas Legalitas Dalam Hukum Pidana Nasional Dan Hukum Pidana Internasional." *Soumatera Law Review* 1, no. 1 (2018): 22–42. <https://doi.org/10.22216/soumlaw.v1i1.3398>.