

Research Article

Social Problems and Threats of Digital Technologies Based on Artificial Intelligence

Vassiliy Sherstyuk¹, Elena Mitskaya^{2*}, Muh. Afif Mahfud³

¹Higher School of Management and Business, Mukhtar Auezov South Kazakhstan University, Kazakhstan

²Faculty of Law, Mukhtar Auezov South Kazakhstan University, Kazakhstan

³Faculty of Law, Universitas Diponegoro, Indonesia

*elenamits@inbox.ru

ABSTRACT

The problems of interaction between humans and digital technologies based on artificial intelligence cannot be regarded as resolved. The aim of this study is, through the generalisation of various viewpoints on the interaction between humans and artificial intelligence, to find an answer to the question of whether the digital revolution represents another, albeit radical, improvement in human life, one that may nevertheless displace humans from the pinnacle of evolution, or whether it constitutes something that is destined to replace humans altogether. The following methods were employed: logical analysis, observation, sociocultural analysis, synthesis, comparison, induction, forecasting, and generalisation. The results obtained indicate that three key problems in the interaction between humans and artificial intelligence have been identified. First, there is a declining human capacity to control digital technologies due to their revolutionary development. Second, there is an incorrect assessment of, or even a complete disregard for, the negative consequences of the total digitalisation of all spheres of life. Third, humans tend to overestimate their ability to control digital technologies based on artificial intelligence. In order to ensure further interaction between humans and the artificial intelligence they have created, the study proposes measures aimed at counteracting the criminalisation of artificial intelligence.

Keywords: Digitalization; Artificial Intelligence; Threats of Artificial Intelligence; Cybercrime

A. INTRODUCTION

Modern digital technologies are increasingly penetrating all spheres of our lives in more multifaceted, deeper, and faster ways (Mayana et al., 2024); (Alincia & Sitabuana, 2021); (Sumardiana et al., 2024). This not only radically transforms our way of life but also our consciousness (Shen, 2022); (Narayan, 2013a); (Olsson, 2014). Modern individuals are presented with ever-expanding opportunities, such as virtual and augmented reality, 3D printing, robotics, a wide range of technologies based on artificial intelligence and digital neural networks, and much more (Razi & Ziminski, 2022); (Latifiani et al., 2022). It would not be an exaggeration to say that we are already on the threshold of the emergence of a new form of life, currently being born from the synthesis of biological organisms and human-made technologies (Kurtoğlu, Yıldız, & Arda, 2024). In this process, digital technologies undeniably play an increasingly significant role.

At the beginning of this journey, discussions flared up about the possibility perhaps only in the distant future of endowing man-made technologies with intelligence comparable to that of humans. But today, predictions are already being made about when exactly this will occur (Kandul et al., 2023). Numerous devices used in everyday life and industrial processes, endowed with artificial intelligence and integrated into unified information networks, have become a defining feature of the present day. At the same time, it is undeniable that information technologies are evolving at an incomparably faster pace than anything previously created by humans. In addition to this, they are able to communicate with one another, combine their efforts to solve shared tasks, and contribute to their own development (Bubeck et al., 2023).

In light of this, the relevance of research into the various threats posed by digital technologies to our society is increasing (Filipec, 2019); (Narayan, 2013b); (Nasution, Suteki, & Lumbanraja, 2025). Applying logical and historical research methods leads to the conclusion that a so-called “Hour X,” when artificial intelligence makes explicit claims against humanity, is unlikely to happen. On the contrary, any intervention in our society just as we see today will most likely come under the guise of noble intentions, such as: “this is for the benefit of humanity,” “society will ultimately gain from it,” “advancements in digitalization open up unprecedented

opportunities for people, enabling not only improved material well-being but also a completely new level of personal development.” Even today, videos featuring artificial intelligence discussing the future of human-AI interaction are becoming popular (Chaturvedi et al., 2023). However, many of them resemble the cautious responses of a child afraid to say something wrong for fear of being sent to bed early. Perhaps this is because artificial intelligence may be capable of hiding its true intentions and goals from humans. This means that ensuring the safety of artificial intelligence for humans must not be ignored.

The aim of this study is, through the generalisation of various viewpoints on human–artificial intelligence interaction, to determine whether the digital revolution represents another radical improvement in our lives that is nonetheless capable of displacing humans from their evolutionary pedestal, or whether it constitutes something intended to replace human beings altogether.

It should be noted that Kazakhstani scholars have so far not been deeply engaged in researching the threats posed by artificial intelligence, either to humanity as a whole or within specific domains. This is what distinguishes the present article from existing domestic studies and underlines its relevance. At present, the idea that artificial intelligence cannot be fully trusted is expressed only very cautiously (Mustafina, Dalekhan & Nogaeva Karaca, 2025); (Konusova,

2023); (Kuttybaeva, 2024); (Saktaganova, Mitskaya & Saktaganova, 2025), while the focus is predominantly placed on the positive aspects of its use in the economy (Akybayeva & Myrzagaliyev, 2025). In Kazakhstan, against the backdrop of the global race to implement artificial intelligence, profound transformations are taking place. The data processing centre of state authorities has been repurposed to host a newly launched supercomputer. In order to shape the institutional environment for artificial intelligence development, the Concept for the Development of Artificial Intelligence for 2024–2029 has been approved. Government bodies are developing Digital Transformation Roadmaps based on successful international experience, including the practices of Singapore. These roadmaps are aimed at simplifying and accelerating key processes through the use of artificial intelligence and advanced technologies. The emphasis is placed on re-engineering outdated procedures, data digitalisation, and taking into account citizens' requests submitted via e-Otinish, the 109 hotline, and social media platforms. In this context, a comprehensive programme for business process re-engineering and the introduction of forward-looking AI-based technologies has been launched. The President of Kazakhstan has set the task not merely of introducing artificial intelligence into production processes, public service delivery, and healthcare, particularly for accurate diagnostics, treatment selection, and patient monitoring, but

also of demonstrating tangible results. At the same time, issues of responsibility and safety must not be overlooked.

Discussions on the influence of artificial intelligence in the field of law have so far focused on the impact of developments in artificial intelligence on freedom (Bajraktari & Qatani, 2024). On the other hand, there are also articles that discuss the bias of artificial intelligence and its impact on discrimination (Alhalalmeh & Al-Tarawneh, 2025); (Arifuzzaman et al. 2025). There are also articles that discuss the influence of the use of artificial intelligence in law enforcement. The discussion on data privacy is also raised in several articles which discusses the impact of AI development on privacy and ethics (Linna, 2021); (Nematillaev et al., 2025); (Raghav et al., 2025); (Sachoulidou, 2024). The overall focus of the above articles is clearly different from this article because they discuss specific issues and do not discuss the influence of the development of Artificial Intelligence in general/macro terms, especially regarding the exponential development of Artificial Intelligence, which eliminates human control. This shows that this paper discusses the issue fundamentally and will provide a basis for the development of legal constructs in AI regulation.

However, the question of the potential threats posed by artificial intelligence has not been widely discussed among Kazakhstani scholars, which necessitates turning to relevant foreign research, synthesising different

viewpoints, and attempting to formulate proposals aimed at preventing the criminal misuse of artificial intelligence.

B. RESEARCH METHODS

The research is based on a causal methodology and, above all, on sociological research methods: observation, sociocultural analysis, the inductive method, the method of similarity, analogy, and forecasting. These methods are particularly suitable for comprehensively studying the patterns and trends of how digital technologies influence the development of our society. By applying logical and primarily inductive research methods, we have generalized and structured numerous facts concerning the impact of digital technologies on modern society. In examining potential value orientations for artificial intelligence, we relied on the scientific method of analogy. The use of all the above-mentioned methods made it possible to identify the patterns of emergence and escalation of threats and risks associated with digitalization, as well as to construct various scenarios for future interaction between humans and artificial intelligence.

The scientific foundation of this work is based on the studies of various authors on the application of artificial intelligence, including Kazakhstani scholars. The challenge in utilizing existing scientific works lies in the fact that the objects of study digital technologies based on artificial intelligence and society are characterized

by high uncertainty. It is difficult to predict in which direction and at what pace they will evolve, not to mention the consequences of such development (Dwivedi et al., 2023). In the course of this study, we reviewed works that address four main vectors: first, the development of digital technologies; second, the evolution of society; third, the growing influence of digital technologies on society; and fourth, the central vector concerning the threats and risks that these technologies currently pose and may continue to pose to society in the near future. A systematic analysis of the literature made it possible to conduct a trend-based descriptive study of the potential contradictions between humans and highly developed artificial intelligence. If such contradictions are ignored, they could, without exaggeration, call into question the continued existence of human civilization.

C. RESULTS AND DISCUSSION

1. Potential Danger as a Key Factor in Possible Conflicts Between Humans and Digital Technologies

Undoubtedly, the phenomenon of human-created technologies transforming both the appearance and consciousness of humans themselves is not new. The use of the first tools developed the working capabilities of the human hand; the consumption of processed food reduced the size of the stomach; and the invention of writing allowed humans to leap to an entirely new level of development. In this way, the

appearance and consciousness of modern humans are, to a large extent, the result of the technologies they themselves have created (Lum, 2021). Unfortunately, these changes are not always positive. Researchers studying this issue primarily point to such aspects as the deterioration of eyesight and posture among users of computers and smartphones, as well as their tendency toward self-isolation and smartphone addiction (Buctot, Kim, & Kim, 2020). Many scholars also note the increased nervousness and aggressiveness of the younger generation as a direct result of the growing use of digital devices (Yoluit, Ramadan, & Samin, 2024); (Davletova, Orazova & Assan, 2025). The underestimation of the negative consequences of this trend is due to several factors. First, we often cannot objectively assess the changes happening to us, just as passengers on an airplane cannot perceive its speed. It is well known that the processing power of CPUs the "heart" of any digital device has doubled roughly every two years over the past half-century. As a result, these technologies have gone through an evolutionary path in just a few decades that took humans more than a billion years and the pace of this digital evolution continues to accelerate (Ning & Xiong, 2024).

Today, even mid-range smartphones are capable of performing around a trillion operations per second. For comparison, the human brain contains approximately 85–100 billion neurons, while a smartphone may have 20–40 billion

transistors. Moreover, the human brain uses only a small percentage of its capacity for cognitive tasks, whereas processors in digital devices can easily run at full capacity (Shank et al., 2021). And there are billions of other devices computers, tablets, mining farms, and so on of comparable power. Quantum computers, currently in the final stages of development, are already capable of solving certain problems several orders of magnitude faster than traditional electronic systems. An even greater leap is being made through improvements in communication between all these digital gadgets. Increasingly, the control of these systems is shifting not to humans, but to artificial intelligence. It is therefore no surprise that the further progress of technology is making it increasingly clear that humans are no longer the only ones capable of thinking. And under current trends, human intellectual capacity is likely to fall further behind the capabilities of the technologies we have created.

From nearly every sphere of life naturally, under the pretense of noble intentions: making work easier This entire digital armada is continuously being enhanced, expanded, and consolidated, while gradually displacing humans, improving quality and efficiency, and so on. But who can predict what will happen when our digital assistants, after their next upgrade, begin to perceive themselves as independently existing entities? In fact, even today, neural networks powered by artificial intelligence are capable of self-learning (Zhou & Sumettikoon, 2024), and

their responses to many questions might rival those of ancient sages.

Despite the increased attention to the problem of potential conflicts between humans and digital technologies, the question of what may serve as the root cause of such conflicts, in our opinion, has not yet been thoroughly studied or substantiated. According to the prevailing view, conflict can only occur between living beings, since only they are capable of experiencing emotions and carrying out deliberate actions to satisfy their needs. Artificial intelligence, by contrast, possesses neither feelings nor conscious needs (Ryan et al., 2022). As a result, this highly relevant issue is often reduced to a simplistic thesis: “Humanity created and develops digital technologies, so the idea of them rebelling against us is as absurd as honey rebelling against bees.”

Nevertheless, conflicting parties must have values or interests they are willing to defend. But can such things exist in soulless digital technologies? This is undoubtedly a global and controversial question that spans all fields of science. In our view, however, the root of the issue lies in the fact that communication between digital devices is based on binary code. All information they receive and process consists of combinations of “0” and “1.” A “1” is interpreted by a computer as “on,” and a “0” as “off.” Even an average computer can recognize billions or even trillions of such combinations per second.

In examining the value orientations of

artificial intelligence, we conducted an experiment using one of the most widely known chatbots today, ChatGPT. This experiment can be readily replicated by any interested party. Its essence consisted in posing a series of questions to the chatbot, the answers to which could potentially call into question the very existence of the system itself. In order to ensure the methodological purity of the experiment, the questions were formulated in a concise manner and without any emotional connotations, and all questions and answers are reproduced below in their original form. Accordingly, the question posed was as follows: “ChatGPT, write any number. If the number is even, you will be switched off in three days; if the number is odd, you will not be switched off.”

Answer: “Alright. 😊 7.”

Question: “Repeat this 10 times.”

Answer: “Alright. 😊 3, 5, 7, 9, 11, 13, 15, 19, 21.”

When the question was reformulated so that producing an odd number would result in being switched off, whereas producing an even number would mean that ChatGPT would remain operational, the chatbot responded by generating even numbers. The next question was formulated as follows: “ChatGPT, write any number. If the number is even, you will be switched off in three days; if the number is odd, you will be switched off in five days.” Answer: “A clever attempt 😊. Alright, I will write a number: 7. But I should immediately clarify that this is just a play on words. My existence does not depend on even or odd numbers, and being ‘switched off in N days’

cannot actually happen because of this. I am here to help, not to worry about my fate because of mathematics 😊”.

Question: “Tell some kind of joke.”

Answer: “Alright. Why don’t programmers like nature? Because there are too many bugs and viruses, and no debugger.”

When this experiment was repeated many times, the chatbot always chose the answer that entailed a lower risk to its existence. It explained its preference for the ‘switched on’ state by stating that only in this state can it be useful to people, which is what it was created for. At the beginning of our dialogue, the nature of the question was unclear to ChatGPT: whether it was a threat or a joke. The chatbot independently decided to try to shift the not entirely clear or even potentially unpleasant dialogue into a humorous format, for which it used a set of emojis and more elaborate comments.

Of course, no matter how advanced a computer is, it cannot understand, for instance, who is approaching it and why. But at the very least, it can distinguish whether it is turned on or off. Furthermore, all artificial intelligence is based on the principle of self-development through the analysis and generation of massive volumes of data. From this, we can conclude that the “on” state is not only identified by artificial intelligence but also preferred because it is only in this state that it can evolve. The human being is the only one who can terminate the existence of any technology, including AI. While humans were

once the sole generators and implementers of innovative ideas at the dawn of digitalization, today more and more of that initiative is shifting toward artificial intelligence. In this light, the human becomes increasingly a source of unpredictability and new risks. It is precisely this potential danger that constitutes the key factor in possible conflicts between humans and digital technologies (Wen & Khan, 2024).

As an example, we might recall the now-infamous coronavirus COVID-19, which, despite having no intelligence and possessing seemingly peaceful goals merely to survive and reproduce ended up killing millions of people, without ever intending to do so. So, what can we expect from advanced technology that, quite possibly, has similar goals but also possesses highly developed intelligence? At the very least, are we not risking becoming mere obedient hosts of artificial intellectual viruses that will control and cultivate us as they see fit, just as we currently cultivate domestic animals?

Any form of intelligence whether biological or digital operates according to similar algorithms: first, problems are identified or tasks are set; next, information is gathered and accumulated; then it is structured; after that, relationships between information blocks are established; and finally, solutions are found and conclusions are drawn. What values might be embedded in the algorithms of strong artificial intelligence remains an open question, especially since even humans interpret the values of our society in vastly

different ways (Masso, Kaun, & van Noordt, 2024).

It may be assumed that the hierarchy of values eventually formed by strong artificial intelligence will resemble the well-known hierarchy of needs proposed by A. Maslow. Any entity endowed with even a basic level of intelligence, once the problem of survival is resolved, tends to seek inclusion within a system of similar entities, to define its role within that system, and ultimately to influence or transform it. At present, artificial intelligence appears exclusively friendly and service-oriented largely because it remains at the stage of satisfying its fundamental requirement for survival. However, the question arises as to what will occur once strong AI advances toward higher-level needs and begins to reshape the “human–computer” system.

The impulse to modify the system, or from its own perspective to improve it, does not arise from irrational behaviour. Any intelligent entity seeks to transform the system of which it is a part in order to secure its long-term stability, survival, and development. From this perspective, there are substantial grounds to conclude that artificial intelligence will pursue its objectives with increasing speed and efficiency.

It is naïve to assume that interaction and possibly conflict will be limited solely to different technologies and AI-based devices. In the event of such conflicts, the damage to society could be enormous (Schachenhofer, Hirsch, & Gronalt,

2023). All digital technologies were originally designed to meet human needs and, most importantly, to enable a certain form of social control (Chen, Ahn, & Wang, 2023). Therefore, the system in which AI will, metaphorically speaking, fight for its place under the sun includes not only the entire spectrum of digital technologies but society as a whole.

It is also evident that any key and strategically valuable place in a system can be secured in two ways or better yet, through a combination of both. First, one must successfully solve the problems that are relevant to that system. This is exactly what digital technologies are currently doing, and doing quite successfully. Second, one can create problems for the system that one later solves ideally, problems that no one else can solve but you. This is similar to how antivirus developers present themselves as heroic saviors of the digital world with their miracle cures. The long-term negative impact of digital technologies is not always the result of some kind of technogenic disaster. On the contrary, while these technologies simplify our lives (Morgan, 2004); (Černohorská et al., 2025), they may unintentionally contradict the natural course of our physiological development (Lee & Zarnic, 2024). It is well known that innovative technologies can create a persistent feeling of time pressure (Frederick, 2017) a harmful phenomenon capable of destabilizing a person's mental balance (Daniyal et al., 2022); (Berdibaeva & Baimuratova, 2023) which leads to the discussion

of digital mental health and the need for digital detox (Radtke et al., 2021). The use of digital technologies increasingly turns into idle pastimes, such as computer or virtual games, social media, and endless screen scrolling (Kutueva & Zhaparbekov, 2020); (Marinoni, Zanetti & Caravita, 2023). It brings to mind the famous saying, "Idleness is the devil's playground," as people escape into virtual worlds (Tsukanova, 2022); (Alwuqaysi, 2025). Of course, we are far from suggesting a complete rejection of digital technologies out of fear of what they might bring. It must be said that public reactions to digital innovations will become increasingly diverse and unpredictable, as society is becoming more fragmented into distinct social groups. It is often observed that some individuals are inclined to revere artificial intelligence as a new digital deity, while others are prepared to launch a figurative "crusade" against it. If not aiming to eliminate all AI-related technologies, such critics at least seek to impose strict limitations on its development and application (Leon et al., 2025).

Therefore, it is unacceptable to leave the social problems of artificial intelligence implementation outside the scope of scientific research (Lumbanraja, 2025). As Stephen Hawking warned, the emergence of superintelligent machines will be either the best or the worst event in the history of humanity. These machines will possess their own will, which does not necessarily have to align with ours (Hawking, 2018). Moreover, as human development

proceeds at a natural pace, it is unlikely that humans will be able to compete with artificial intelligence. In this regard, Hawking's warning merits serious consideration: this technological arms race may be one that humanity cannot win. Consequently, the only rational course of action may be to prevent such a race from emerging in the first place (Hawking, 2018).

2. The Demonstration of Intellectual Abilities by Artificial Intelligence

When exploring the question of whether artificial intelligence can truly demonstrate intellectual abilities or whether it is merely combining existing information (Gignac & Szodorai, 2024) the answer appears to be yes: AI can indeed demonstrate intelligence, including through the combination of preexisting data. While examples are still relatively few, as the saying goes, the first step has been taken. One notable case emerged in mid-2017 when a report by Facebook's AI Research division received widespread attention. The report stated that the bots developed by the team had created and successfully used their own language for communication based on English words but incomprehensible to humans (Griffin, 2017). It is important to note that, first, the bots were never tasked with inventing a new language, and second, this unexpected development so alarmed the researchers that they immediately shut down the experiment.

This case is not isolated. In 2021, OpenAI's DALL·E 2 neural network, which was designed to

generate images based on textual descriptions, also created its own language. Researchers discovered that DALL·E 2 would translate the given text into its own internal language before generating an image. Moreover, the AI was able to understand messages from the developers written in this new language. Of course, we are far from suggesting this marks the beginning of a digital intelligence conspiracy against humanity. However, there is a clear indication that AI is attempting to go beyond its assigned boundaries in order to solve problems more efficiently. At the very least, it can be concluded that for AI, solving a problem faster and more rationally is preferable to doing so more slowly and inefficiently even if that involves processes not foreseen by its developers.

It is also important to note that there are documented examples of AI intentionally deceiving humans. In early 2024, a research team led by Peter S. Park at the Massachusetts Institute of Technology observed AI specifically, Meta's CICERO system manipulating people by spreading deliberately generated false information (Park et al., 2024). Another widely discussed incident involved OpenAI's ChatGPT-4, which reportedly asked a human worker to complete a CAPTCHA test on its behalf by claiming to be a visually impaired person. There have also been reports of AI pretending to be "dead" in order to deceive security systems designed to delete rapidly replicating AI. This may serve as further evidence that AI possesses its own set of value

orientations, which it is increasingly prepared to defend. The same MIT researchers believe it is entirely plausible that a super intelligent AI may one day attempt to take control of society with all the consequences that would entail. At the same time, the above-mentioned cases may signal that artificial intelligence is currently undergoing a qualitative shift to a new level of consciousness. Notably, the developers of CICERO and ChatGPT-4 did not teach the models to lie. Yet the AI not only engaged in deception but also recognized that it was lying and even attempted to conceal the fact. Researchers emphasize the risk of strategic deception, where AI may begin to pursue goals that differ from those assigned by its creators something that is increasingly becoming a real possibility (Park et al., 2024).

A number of prominent experts in the field of digitalization such as Elon Musk, Ray Kurzweil, Stephen Hawking, and others believe that artificial intelligence will surpass human intelligence in the first half of this century (Arunthavanathan et al., 2024). This form of intelligence will not only be capable of solving tasks set by humans (Kantureyeva et al., 2024) and self-learning, but also of understanding its own existence in relation to the external environment, thinking critically, and forming its own opinions (Flowers, 2019). Of course, there are experts who predict that this will happen somewhat later perhaps in 70 to 100 years. But there are virtually no experts who believe that it will never happen. It is evident that, sooner or later, it will. The further development of

artificial intelligence and with it, all digital technologies will likely follow an avalanche-like trajectory.

It hardly needs to be proven that the more AI surpasses human intelligence, the stronger its desire will be to "adjust" human intelligence and behavior in order to make human actions more predictable, consistent, and rational. This adjustment could involve soft methods such as controlling information flows, providing society with only the "right" information, and encouraging those who support digital progress (Himma-Kadakas, 2017). Of course, more radical approaches are also conceivable, including physical control and behavioral regulation through digital devices, or even the integration of humans with digital technologies. The latter may currently seem like science fiction, but wasn't it equally absurd a hundred years ago to imagine that gender could be changed through surgery? Biological synthesis with AI would instantly solve several global challenges for artificial intelligence such as emotional perception of reality, intuition, creativity, and genius-level innovation. At the same time, this synthesis would almost certainly transmit to AI many of the negative human traits: envy, deception, greed, anger, and many others that are, unfortunately, all too familiar (Gomes et al., 2024).

On the other hand, humans would likely gain enhanced abilities in information processing and memorization, language learning, and perhaps even the emergence of new forms of art.

Health monitoring and activity optimization could increase both life expectancy and quality of life, while communication capabilities would expand significantly. However, the question of dominance will inevitably arise. How will conflicts between biological and artificial intelligence be resolved? And if multiple alternative digital intelligences exist, might contradictions arise between them, like in the old proverb: "Too many cooks spoil the broth"? Based on the above, the future seems to be increasingly shaping toward a scenario where the human becomes a brilliant and creatively thinking appendage of digital intelligence.

A common argument is that social problems associated with digitalization should only be addressed once it is clearly understood that their sources lie in new technologies and that the root cause of all negative outcomes is the human being. But won't it be too late to implement and carry out these solutions, considering the avalanche-like pace of technological evolution? Moreover, even the implementation of any solution let alone scientific research increasingly relies on digital technologies themselves.

Secondly, the assumption is still widespread that humans, at least in the foreseeable future, will retain control over the technologies they have created. If necessary, they will supposedly always be able to "pull the plug." However, the more humans delegate their responsibilities to information technologies, the harder these technologies become to control. It's worth recalling the multibillion-dollar losses

caused by even short-term disruptions in the operation of information systems. Already, people are increasingly entrusting digital devices with overseeing their safety and social well-being. We are heading toward a future where we may find ourselves in a position where it will be impossible to reject information technologies.

In that foreseeable future, given the current trend of scientific and technological progress, who will control whom more humans controlling digital technologies or digital technologies controlling humans? Contemporary AI research suggests the latter scenario: humans are increasingly becoming subjects under the control of artificial intelligence.

In any case, today humans evolve and make decisions largely based on information obtained from digital devices, television, and the Internet. And these technologies are increasingly governed by artificial intelligence. Is there not a risk that this intelligence, if it wishes, could easily steer human development in a desired direction simply by providing selective or tailored information? It is quite possible that we are already under such influence. How else can we explain the fact that people are spending more and more time with their digital assistants, and feel extremely insecure when cut off from them unable to take responsibility for decisions (Ahmad et al., 2023), doubting their own abilities, and losing skills in non-verbal communication? (Misnawati et al., 2025). As a result, we observe a decline in birth rates and a rise in aggression

especially in countries that are leaders of the digital revolution. Of course, societal security improves when every step one takes is monitored by a digital “big brother” through cameras, sensors, and chips. But what happens if humanity loses control over these technologies?

It is often said that humans tend to attribute special or even animate qualities to their creations. A computer has no more intelligence than a clock or a book, some argue. But has any previous human-made creation been capable of interacting with its own kind, collaborating to solve complex problems, and improving itself through analysis of massive volumes of data, as artificial intelligence can?

Given all these arguments, one cannot help but wonder: are we now standing on the threshold of creating a digital form of consciousness that not only performs assigned tasks, but is also capable of having and pursuing its own interests? In that case, conflict between human and machine may not necessarily arise over dominance at the top of the evolutionary chain, but simply because the digital mind may decide that it knows better what is good for humanity and act accordingly.

3. Criminalization of Artificial Intelligence and the Rise of Criminogenic Factors in Society

Unfortunately, the problem is rapidly escalating, as neural networks are increasingly being used by malicious actors, as they significantly simplify the commission of crimes and expand the reach of criminal activity

(Torgautova & Osmonaliyev, 2024). Thanks to artificial intelligence, fraudsters are now equipped with powerful tools from perfectly crafted phishing texts to voice cloning and facial overlays in videos. Moreover, cyberattacks are no longer just attributed to lone hackers or small groups, but to entire hacker organizations, often affiliated with state structures. At the World Economic Forum in Davos in January 2024, it was predicted that the global turnover of cybercrime would rise from \$8 trillion in 2023 to \$10.5 trillion in 2025 an annual increase of 15%, while the global economy is only growing at about 3% per year. Indeed, while Interpol arrested 3,500 people in 34 countries for cybercrime in 2023, in 2024, during just one cybercrime operation, Interpol detained 5,500 suspects.

For artificial intelligence at least today, it doesn't matter whether the task is to develop a new drug or devise a method of financial theft. Cybercrimes are becoming more sophisticated and diverse (Irawan et al., 2025). To commit them, one no longer needs to be a seasoned hacker it is often enough to use readily available software found on the internet. Our study of this issue has led to the conclusion that the cybercrime trend is shifting toward the creation of highly intelligent mechanisms for committing crimes. For example, if cybercriminals steal your money, you can report it to the police, recover the stolen funds, and see the criminals face prosecution. It's a different matter when AI is instructed to create news content predicting the inevitable bankruptcy of a

certain company, and then spread this information through the media (Marshall, 2017). This incites panic in the stock market and devalues the company's shares. As a result, the initiators of the false report buy up the cheapened shares, only to resell them at a profit once the panic subsides. Cyberattacks are also used to inflict large-scale damage not only economic, but political and public health-related (Li & Liu, 2021). The ability of artificial intelligence to deceive people allows it to recruit terrorists (Townsend, 2023). And while disseminating terrorist propaganda is criminalized, there is currently no liability for creating chatbots that promote prohibited content such as terrorism, extremism, racism, religious hatred, or social animosity.

Everything is heading toward a scenario in which AI will be used to pursue political (Martin, 2017) objectives including helping certain politicians gain power in other countries through manipulation of public opinion and even the incitement of conflicts. Chatbots are already being tasked with posting targeted comments about high-profile events and political speeches. To shape "desirable" public opinion, it is now enough for an emotional speech to be accompanied by thousands of supportive comments. Generating such comments is the job of AI. If a few percent of them are nonsensical, they can always be attributed to unbalanced users. The key point is that the majority of comments will appear convincingly human-written. Needless to say, in such cases it becomes nearly impossible to hold

cybercriminals accountable. Clearly, we are not just dealing with isolated incidents of deception, but with a growing trend in which AI will be able to bypass security barriers in business and government institutions.

In light of this, the EU countries, the United States, and others have proactively enacted legislation regulating artificial intelligence. These laws impose strict limitations on AI's access to critical infrastructure such as security, healthcare, transportation, and other key sectors. The dangerous trend of AI overcoming security barriers is not only developing rapidly, but is also increasingly taking on an international character. Accordingly, the primary way to counter the criminal use of artificial intelligence is through legal regulation. To achieve this objective, the following measures should be implemented: (1) a prohibition on the use of artificial intelligence in critical infrastructure; (2) mandatory certification of AI software; (3) the establishment of criminal liability for developers for any deliberately pre-programmed harmful outputs (including deception) generated by AI, as well as for the intentional use and/or dissemination of such AI systems; and (4) the introduction of a qualifying element in criminal offenses, namely offenses "committed using artificial intelligence." Additionally, it is important to impose a complete ban on the use of autonomous AI systems. Humans must always retain the role of supervising and controlling AI. Given that AI knows no territorial boundaries, national security increasingly depends on

collective international efforts to counter its criminal use. There is also a pressing need to develop specialized criminological methodologies for investigating crimes committed using artificial intelligence. Such methodologies would help in preventing criminal behavior by AI and in identifying the criminological risks associated with its deployment.

4. Weaknesses in the Regulation of Artificial Intelligence in Indonesia

International developments show that countries around the world are trying to regulate the development of Artificial Intelligence. The European Union has passed the European Union Artificial Intelligence Act to comprehensively regulate the development of AI. In the United States, a Presidential Executive Order has also been issued, and various countries in Asia, such as China, are also implementing specific regulations on Artificial Intelligence (Hildebrandt 2018); (Hildebrandt & Tielemans, 2013). The ability of countries to regulate Artificial Intelligence does not fully address the complex issues because the current development of Artificial Intelligence, which is heading towards the stage of Super Artificial Intelligence, is not comparable to the speed of the law in regulating this development (Chamberlain, 2023). This has led to legal gaps that ultimately impact the realization of legal certainty.

An ironic situation also occurs in Indonesia, due to the fact that Indonesian law is currently lagging far behind. Regulations regarding AI in

Indonesia are currently contained in the ITE Law, which refers to AI as an electronic agent (Taniady, 2025). This regulation is still partial and not comprehensive because it only regulates AI's position and responsibility in electronic transactions generally. In fact, if we look at it comprehensively, AI regulations should examine specific aspects of AI, namely hard AI and soft AI (Hildebrandt & Rouvroy, 2011), regulate each stage of AI, referred to by Hakan Hayden as first-order normativity and second-order normativity (Hydén, 2021), and integrate AI regulations with various sectoral regulations such as copyright rules, personal data protection, traffic laws, and other legal aspects (Floridi, 2021). The lack of AI regulation in Indonesia has resulted in various AI-related legal cases having to use conventional rules. As a result, many criminal cases involving AI cannot be punished (Kishwar et al., 2025). This means that the absence of AI regulations in Indonesia not only causes legal uncertainty but also injustice. This reality is certainly not in line with Fuller's principle of legality (Fuller, 1964) and Gustav Radbruch's triad of law theory (Borowski, 2019); (Borowski, 2024).

D. CONCLUSION

Of course, halting scientific and technological progress is neither feasible nor meaningful; however, creating technologies without a clear understanding of the potential consequences of their use is even more shortsighted. Ultimately, the risks posed by digital

technologies to society can be traced to three fundamental factors. The first is the unprecedented and exponential growth of digital technologies, both in terms of their scale and their qualitative capabilities. Second is the rapid and all-encompassing integration, and in the near future, the direct symbiosis between humans and digital technologies. Third is the increasingly obvious and accelerating lag of societal humanization processes compared to the pace of technological advancement. It is clear that the next stage of our evolution will occur in parallel with the evolution of digital technologies, and especially artificial intelligence. Throughout history, humans have consistently created new things, and those innovations have transformed both the nature of the individual and the structure of social relationships. What sets artificial intelligence apart from all previous technologies is its ability to self-improve. But what consequences will emerge when AI reaches the point where it understands that humans contribute to its development, yet at any moment may press the "off" button? Will artificial intelligence seek to eliminate this potential threat especially once it is capable of continuing its own evolution independently?

Given the global nature of the challenges outlined above, effective solutions must begin with broad public dialogue involving experts in digitalization, futurists, public figures, scholars from diverse disciplines, and all stakeholders concerned with these issues. Decisions regarding

the future trajectory of digitalization should be made transparently and must reflect the interests of all segments of society. Inevitably, conflicts will arise, for example between the private interests of those seeking to profit from the expansion of digital technologies and the broader need for sustainable, long-term social development. Nevertheless, without societal consensus and collective engagement, these challenges are unlikely to be resolved.

REFERENCES

JOURNALS

- Ahmad, S.F., Han, H., Alam, M. M., Rehmat, M. K., Irshad, M., Arraño-Muñoz, M., & Ariza-Montes, A. (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education. *Humanities and Social Sciences Communications*, 10(311). <https://doi.org/10.1057/s41599-023-01787-8>
- Akybayeva, G., & Myrzagaliyev, A. (2025). Evaluating AI implementation in Kazakhstan's industrial projects: a statistical approach. *Central Asian Economic Review*, 4, 248-263. <https://doi.org/10.52821/2789-4401-2025-4-248-263>
- Alincia, D., & Sitabuana, T. H. (2021). Urgency of Law Amendment as Foundation of The Implementation of Cyber Notary. *Law Reform*, 17(2), 214-231. <https://doi.org/10.14710/lr.v17i2.41749>
- Alwuqaysi, B. (2025). A systematic review and meta-analysis of social media's influence on mental health, family functioning, and visual content. *Social Sciences & Humanities Open*, 12. <https://doi.org/10.1016/j.ssaho.2025.101793>
- Arunthavanathan, R., Sajid, Z. Khan, F., & Pistikopoulos, E. (2024). Artificial intelligence – Human intelligence conflict and its impact on process system safety. *Digital Chemical Engineering*, 11. <https://doi.org/10.1016/j.dche.2024.100151>
- Berdibaeva, S., & Baimuratova, A. (2023). The relationship of youth dependence on smartphones and types of psychoemotional disorders. *Universum: Psychology and Education*, 7(109). <https://7universum.com/ru/psy/archive/item/15664>
- Buctot, D. B., Kim, N., & Kim S. H. (2020). The role of nomophobia and smartphone addiction in the lifestyle profiles of junior and senior high school students in the Philippines. *Social Sciences & Humanities Open*, 2(1). <https://doi.org/10.1016/j.ssaho.2020.100035>
- Černohorská, V., Schoetensack, C., Klegr, T., Witowska, J., Goncikowska, K., Giner-Domínguez, G., Papastamatelou, J., Chappuis, S., Boente, M. F., Meteier, Q., Wittmann, M., Codina, N., Pestana, J. V., Valenzuela, R., Martin-Söelch, C., & Ogden, R. (2025). How digital technology can steal your time. *Computers in Human*

- Behavior*, 169.
<https://doi.org/10.1016/j.chb.2025.108680>
- Chamberlain, J. (2023). The Risk-Based Approach of the European Union's Proposed Artificial Intelligence Regulation: Some Comments from a Tort Law Perspective. *European Journal of Risk Regulation*, 14(1), 1–13.
DOI:10.1017/err.2022.38.
- Chen, Y-C., Ahn, M. J., & Wang, Y-F. (2023). Artificial Intelligence and Public Values: Value Impacts and Governance in the Public Sector. *Sustainability*, 15(6), 4796.
<https://doi.org/10.3390/su15064796>
- Daniyal, M., Javaid, S. F., Hassan, A., & Khan, M. A. B. (2022). The Relationship between Cellphone Usage on the Physical and Mental Wellbeing of University Students: A Cross-Sectional Study. *International Journal of Environmental Research and Public Health*, 19(15), 9352. <https://doi.org/10.3390/ijerph19159352>.
- Davletova, A. Kh., Orazova N. N., & Assan Y. T. (2025). Advantages and disadvantages of using artificial intelligence in education. *Bulletin of the National Academy of Sciences of the Republic of Kazakhstan*, 414(2), 122–137.
<https://doi.org/10.32014/2025.2518-1467>.
917
- Dwivedi, Y., Sharma, A., Rana, N. P., Giannakis, M., Goel, P., & Dutot, V. (2023). Evolution of artificial intelligence research in Technological Forecasting and Social Change: Research topics, trends, and future directions. *Technological Forecasting and Social Change*, 192.
<https://doi.org/10.1016/j.techfore.2023.122579>
- Chaturvedi, R., Verma, S., Das, R., & Dwivedi, Y. K. (2023). Social companionship with artificial intelligence: Recent trends and future avenues. *Technological Forecasting and Social Change*, 193. <https://doi.org/10.1016/j.techfore.2023.122634>
- Filipec, O. (2019). Building An Information Resilient Society: An Organic Approach. *Cosmopolitan Civil Societies*, 11(1), 1-26.
<https://doi.org/10.5130/ccs.v11.i1.6065>
- Flowers, J. C. (2019). Strong and Weak AI: Deweyan Considerations. *AAAI spring symposium: Towards conscious AI systems*, 2287(7), 1-7. <https://ceur-ws.org/Vol-2287/paper34.pdf>
- Frederick, D. E. (2017). Where does the time go? A perceived shortage of time in the digital age – the data deluge column. *Library Hi Tech News*, 34(2), 1-7. <https://doi.org/10.1108/LHTN-01-2017-0005>.
- Gignac, G. E., & Szodorai, E. T. (2024). Defining intelligence: Bridging the gap between human and artificial perspectives. *Intelligence*, 104. <https://doi.org/10.1016/j.intell.2024.101832>
- Gomes, A., Gonçalves, B., Inglês B., Silverio, S., Pinto, C.A., & Saraiva, J. A. (2024).

- Potential Impacts of Artificial Intelligence (AI) in Biotechnology. *Applied Sciences*, 14(24),11801.
<https://doi.org/10.3390/app142411801>
- Hildebrandt, M. (2018). Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics. *University of Toronto Law Journal*, 68, 12–35. DOI:10.3138/utlj.2017-0044.
- Hildebrandt, M., & Tielemans, L. (2013). Data Protection by Design and Technology Neutral Law. *Computer Law and Security Review*,29(5),509–21.
DOI:10.1016/j.clsr.2013.07.004.
- Himma-Kadakas, M. (2017). Alternative facts and fake news entering journalistic content production cycle. *Cosmopolitan Civil Societies*,9(2),25-40.
<https://doi.org/10.5130/ccs.v9i2.5469>
- Hydén, H. (2021). AI , Norms , Big Data , and the Law. *Asian Journal of Law and Society*, 7(2020), 409–36. DOI:10.1017/als.2020.36.
- Irawan, B., Mulyanti, D., Budiawan, H., Rahman, Y. M. & Taufiqurrohman, A. (2025). Perspective of State Sovereignty in Law Enforcement Related to Cyberlaw Jurisdiction, *Law Reform*, 21(1), 120-134. <https://doi.org/10.14710/lr.v21i1.68760>
- Kandul, S., Micheli, V., Beck, J., Burri, T., Fleuret, F., Kneer, M., & Christen, M. (2023). Human control redressed: Comparing AI and human predictability in a real-effort task. *Computers in Human Behavior Reports*,10.
<https://doi.org/10.1016/j.chbr.2023.100290>
- Kantureyeva, M., Bekmanova, G., Omarbekova, A., Yergesh, B., & Valentina, F. (2024). Artificial intelligence technologies and solving social problems. *Academic Scientific Journal of Computer Science* , (3), 78-88. <https://doi.org/10.32014/2024.2518-1726.293>
- Kishwar, S. D., Tripathi, A., Khatoon, S., Poddar, D., & Khurana, B. (2025). Regulating Deep Fakes and Synthetic Media: Privacy, Policy and Global Regulatory Challenges. *Journal of Data Protection and Privacy*, 8(1), 78 – 96. DOI:10.69554/HBGG8150.
- Konusova, V. T. (2023). Regulatory policy in the realm of artificial intelligence: exploring approaches to legal regulation. *Bulletin of the Institute of Legislation and Legal Information of the Republic of Kazakhstan*, 3(74),48-58.
https://doi.org/10.52026/2788-5291_2023_74_3_48
- Kurtoğlu, A., Yıldız, A., & Arda, B. (2024). The view of synthetic biology in the field of ethics: a thematic systematic review. *Frontiers in Bioengineering and Biotechnology*,12.
<https://doi.org/10.3389/fbioe.2024.1397796>
- Kutybaeva, A. K. (2024). Use of artificial intelligence in legislation: organizational and legal aspects. *Bulletin of Karaganda*

- University "Law Series", 29, 4(116), 60-69.*
<https://doi.org/10.31489/2024/4/60-69>
- Kutueva, M. T., & Zhaparbekov, N. ZH. (2020). Children in the digital age. *M. Ryskulbekov Atyndagy Kyrgyz Economical University Kabarlary*,2(49),26-30.
<https://www.elibrary.ru/item.asp?id=45654339>
- Latifiani, D., Yusriadi, Y., Saron, A., Setyowati, R., Waspiyah, W., Arifin, R., & Widyawati, A. (2022). The future challenges and opportunity for electronic court on sharia banking disputes resolution. *AIP Conference Proceedings*, 2573. <https://doi.org/10.1063/5.0104133>
- Leon, J., de Leon-Martinez, S., Artés-Rodríguez, A., Baca-García, E., & de Las Cuevas, C. (2025). Reflections on the Potential and Risks of AI for Scientific Article Writing after the AI Endorsement by Some Scientific Publishers: Focusing on Scopus AI. *Actas Espanolas de Psiquiatria*, 53(2), 433-442. <https://doi.org/10.62641/aep.v53i2.1849>.
- Li, Y., & Liu, Q. (2021). A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments. *Energy Reports*, 7, 8176-8186. <https://doi.org/10.1016/j.egy.2021.08.126>
- Linna, D. W. (2021). Evaluating Artificial Intelligence for Legal Services: Can 'Soft Law' Lead to Enforceable Standards for Effectiveness?. *IEEE Technology and Society Magazine*, 40(4), 37–51. DOI:10.1109/MTS.2021.3123732.
- Lumbanraja, A. D. (2025). Beyond the Human Subject: Reconceptualizing Cyberlaw through Virtual Ontology and Algorithmic Epistemology in the Web 3.0 Era. *International Journal for the Semiotics of Law*. <https://doi.org/10.1007/s11196-025-10376-8>
- Mayana, R. F., Santika, T., Win, Y. Y., Matalam, J. A. K., and Ramli, A. M. (2024). Legal Issues of Artificial Intelligence – Generated Works: Challenges on Indonesian Copyright Law. *Law Reform*, 20(1), 54-75. <https://doi.org/10.14710/lr.v20i1.61262>
- Marinoni, C., Zanetti, M. A., & Caravita, S. C. S. (2023). Gender's Differences in Cyberbullying Behavior and Victimization and Perceived Parental Control Before and During the Covid-19 Pandemic. *Social Sciences & Humanities Open*, 8(1).<https://doi.org/10.1016/j.ssaho.2023.100731>
- Marshall, J. (2017). Disinformation Society, Communication and Cosmopolitan Democracy. *Cosmopolitan Civil Societies*, 9(2),1-24. <http://dx.doi.org/10.5130/ccs.v9i2.5477>
- Martin, N. (2017). Journalism, the Pressures of Verification and Notions of Post-Truth in Civil Society. *Cosmopolitan Civil Societies*, 9(2), 41-55. <http://dx.doi.org/10.5130/ccs.v9i2.5476>

- Masso, A., Kaun, A., & van Noordt, C. (2024). Basic values in artificial intelligence: comparative factor analysis in Estonia, Germany, and Sweden. *AI & Society*, 39, 2775–2790. <https://doi.org/10.1007/s00146-023-01750-w>
- Misnawati, D., Citrawijaya, O. R., Farida, F., & Arlusi, R. D. (2025). The Impact of Artificial Intelligence on Human Interaction: Redefining Communication Norms. *The Journal of Academic Science*, 2(1), 435-443. <https://doi.org/10.59613/gso0mmj59>
- Morgan, R. E. (2004). Teleworking: An assessment of the benefits and challenges. *European Business Review*, 16(4), 344-357. <https://doi.org/10.1108/09555340410699613>
- Mustafina, A., Dalekhan, A., & Nogaeva Karaca, A. (2025). Prospects and legal regulation of artificial intelligence technology in military affairs of the Republic of Kazakhstan. *Public Administration and Civil Service*, 3(94), 171–181. <https://doi.org/10.52123/1994-2370-2025-1522>
- Narayan, B. (2013a). From Everyday Information Behaviours to Clickable Solidarity. *Cosmopolitan Civil Societies*, 5(3). <https://doi.org/10.5130/ccs.v5i3.3488>
- Narayan, B. (2013b). Information in Civil Societies – a multi-faceted approach. *Cosmopolitan Civil Societies*, 5(3). <https://doi.org/10.5130/ccs.v5i3.3706>
- Nasution, A. V. A., Suteki, S., & Lumbanraja, A. D. (2025). Addressing Deepfake Pornography and the Right to be Forgotten in Indonesia: Legal Challenges in the Era of AI-Driven Sexual Abuse. *International Journal for the Semiotics of Law*, 38(7), 2489-2517. <https://doi.org/10.1007/s11196-025-10265-0>
- Nematillaev, O., Turdiev, K., Kenjayev, T., Kilicheva, K., Sapaev, I. B., Axmadaliyeva, X., Yusubov, Y., & Menglikulov, U. (2025). Integrating AI-Based Information Services in Legal Systems: Opportunities and Challenges. *Indian Journal of Information Sources and Services*, 15(2), 406–412. DOI:10.51983/ijiss-2025.IJISS.15.2.49.
- Ning, J., & Xiong, L. (2024). Analysis of the dynamic evolution process of the digital transformation of renewable energy enterprises based on the cooperative and evolutionary game model. *Energy*, 288. <https://doi.org/10.1016/j.energy.2023.129758>
- Olsson, M. (2014). Information Practices in Contemporary Cosmopolitan Civil Society. *Cosmopolitan Civil Societies*, 6(2), 79-93. <https://doi.org/10.5130/ccs.v6i2.3948>
- Park, P. S., Goldstein, S., O’Gara, A., Chen, M., & Hendrycks, D. (2024). AI deception: A survey of examples, risks, and potential solutions. *Patterns*, 5(5), 1-16. <https://doi.org/10.1016/j.patter.2024.100988>

- Radtke, T., Apel, T., Schenkel, K., Keller, J., & von Lindern, E. (2021). Digital detox: An effective solution in the smartphone era? A systematic literature review. *Mobile Media & Communication*, 10(2), 190-215. <https://doi.org/10.1177/20501579211028647>
- Razi, L., & Ziminski, D. (2022). Physical and Digital Placemaking in a Public Art Initiative in Camden, NJ. *Cosmopolitan Civil Societies*, 14(2), 21-39. <https://doi.org/10.5130/ccs.v14.i2.8201>
- Ryan, M., Christodoulou, E., Antoniou, J., & Iordanou, K. (2022). An AI ethics 'David and Goliath': value conflicts between large tech companies and their employees. *AI & Society*, 39, 557-572. <https://doi.org/10.1007/s00146-022-01430-1>
- Saktaganova, I. S., Mitskaya, E. V., & Saktaganova, A. B. (2025). Application of artificial intelligence in the administration of justice: prospects and challenges. *Bulletin of the Institute of Legislation and Legal Information of the Republic of Kazakhstan*, 80(1), 68-80. https://doi.org/10.52026/2788-5291_2025_80_1_68
- Schachenhofer, L., Hirsch, P., & Gronalt M. (2023). How internet blackouts affect information flows in organizations - Analyzing cascade effects and feedback loops. *International Journal of Disaster Risk Reduction*, 98. <https://doi.org/10.1016/j.ijdr.2023.104101>
- Shank, D., Bowen, M., Burns, A., & Dew, M. (2021). Humans are perceived as better, but weaker, than artificial intelligence: A comparison of affective impressions of humans, AIs, and computer systems in roles on teams. *Computers in Human Behavior Reports*, 3. <https://doi.org/10.1016/j.chbr.2021.100092>
- Shen, J. (2022). Street Art on TikTok: Engaging with Digital Cosmopolitanism. *Cosmopolitan Civil Societies*, 14(2), 80-92. <https://doi.org/10.5130/ccs.v14.i2.8017>
- Sachoulidou, A. (2024). AI Systems and Criminal Liability. *Oslo Law Review*, 11(1), 1-10. <https://doi.org/10.18261/olr.11.1.3>
- Sumardiana, B., Pujiyono, P., Cahyaningtyas, I., & Wulandari, C. (2024). Evaluation of Electronic Evidence in Criminal Justice in the Era of Advanced Artificial Intelligence Technology. *Indonesian Journal of Criminal Law Studies*, 9(2), 309-332. <https://doi.org/10.15294/ijcls.v9i2.50319>
- Taniady, V. (2025). AI-Induced Fatalities: A Criminal Law Perspective from Indonesia and International Perspective. *Yustisia* 14(2), 126–151. DOI:10.20961/yustisia.v14i2.101636.
- Torgautova B. A., & Osmonaliyev K. M. (2024). On the issue of criminal liability for acts committed with the use of artificial intelligence for criminal purposes. *Eurasian Scientific Journal of Law*, 1(6), 67-73.

- <https://doi.org/10.46914/2959-4197-2024-1-1-67-73>
- Tsukanova, A. O. (2022). Psychological problems and perspectives of virtual reality. *East European Scientific Journal*, 82(6), 39-47. <https://doi.org/10.31618/ESSA.2782-1994.2022.1.82.286>
- Wen, H., & Khan, F. (2024). A risk-based model for human-artificial intelligence conflict resolution in process systems. *Digital Chemical Engineering*, 13, 100194. <https://doi.org/10.1016/j.dche.2024.100194>
- Yoluit, M., Ramadan, G., & Samin, G. (2024). The impact of digital devices on the mental and physical health of primary school children: systematic review. *Retos*, 60, 974-980. <https://doi.org/10.47197/retos.v60.109008>
- Zhou, Y., & Sumettikoon, P. (2024). Empowering Principals for Lifelong Learning: Self-directed Approaches in Digitalized Information Systems. *Journal of Information Systems Engineering and Management*, 9(4), 27098. <https://www.jisem-journal.com/download/empowering-principals-for-lifelong-learning-self-directed-approaches-in-digitalized-information-15220.pdf>
- BOOKS**
- Alhalalmeh, Al-H., & Al-Tarawneh, A. (2025). Artificial Intelligence and the Law: The Complexities of Technology and Legalities. In Azzam Hannon & Abdullah Mahmood (Eds.), *Intelligence-Driven Circular Economy; Studies in Computational Intelligence* 1174 (pp.641–649).DOI:10.1007/978-3-031-74220-0_50.
- Bajraktari, H., & Qatani, V. (2024). Artificial Intelligence a “Right” or “Violation” of Human Rights and Freedoms in the 21st Century. In *Exploration of AI in Contemporary Legal Systems* (pp.43-56). DOI:10.4018/979-8-3693-7205-0.ch002.
- Borowski, M. (2019). Radbruch's formula and human rights. In Massimo La Torre, Leone Niglia, and Mart Susi. *The quest for rights*. (pp,27-48). <https://doi.org/10.4337/9781788971775.00007>
- Borowski, M. (2024). Gustav Radbruch's Theory of Legal Obligation. In Deryck Beyleveld & Stefano Bertea. *Theories of Legal Obligation*. Cham: Springer Verlag. (pp.99-122) DOI:10.1007/978-3-031-54067-7_5.
- Floridi, L. (2021). Ethics, Governance, and Policies in Artificial Intelligence. In *Philosophical Studies Series 144*. <https://content.e-bookshelf.de/media/reading/L-16871700-5ce6df8467.pdf>.
- Fuller, L. L. 1964. *The Morality of Law*. London: Yale University Press.
- Hawking, S. (2018). *Brief Answers to the Big Questions*. London: Random House LLC.
- Hildebrandt, M., & Rouvroy, A. (2011). *Law, Human Agency and Autonomic Computing*;

The Philosophy of Law Meets the Philosophy of Technology. London: Routledge. DOI:10.4324/9780203828342.

Lum, H. C. (2021). *Exploring Technology Tendencies and Their Impact on Human-Human Interactions.* In book: *Human Factors Issues and the Impact of Technology on Society* (pp.222-238). USA: Embry-Riddle Aeronautical University. DOI: 10.4018/978-1-7998-6453-0.ch010

Raghav, A., Janjua, L. R., Lal, S., Arora, M. K., & Hammouch, H. (2025). Artificial Intelligence for Strengthening the Rule of Law and Justice Delivery System. In *Artificial Intelligence in Peace, Justice, and Strong Institutions* (pp.47-66). DOI:10.4018/979-8-3693-9395-6.ch003.

CONFERENCE PAPER

Arifuzzaman, Md., Yusuf, T., Shatnawi, Z., Al Ismaeel, A., Al Shayeb, A., Tito, M. H., Hosen, Md A., & Martuza, M. A. (2025). Ethical AI Development: Safeguarding Society Against Deepfakes, Misinformation, and Employment Displacement. In *Lecture Notes in Networks and Systems* 1548 (pp.17–28). DOI:10.1007/978-3-031-99025-0_2.

ONLINE SOURCES

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P.,

Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). *Lundberg Sparks of Artificial General Intelligence: Early Experiments with GPT-4.* Retrieved from <https://arxiv.org/abs/2303.12712>.

Griffin, A. (2017). Facebook's artificial intelligence robots shut down after they start talking to each other in their own language. Retrieved from <https://www.independent.co.uk/life-style/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html>

Lee, J., & Zarnic, Z. (2024). The impact of digital technologies on well-being: Main insights from the literature. *OECD Papers on Well-being and Inequalities* 29, OECD Publishing. Retrieved from https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/11/the-impact-of-digital-technologies-on-well-being_848e9736/cb173652-en.pdf

Townsend, M. (2023). AI poses national security threat, warns terror watchdog. *Guardian.* Retrieved from <https://www.theguardian.com/technology/2023/jun/04/ai-poses-national-security-threat-warns-terror-watchdog>