

<http://doi.org/10.22133/mtlj.2025.461599.1334>

The Requirements of Liability of Artificial Intelligence in the Light of Criminal and Civil Law

Yasaman Khademi Shoormasti^{1*} , Mohammad Yekrangi² 

¹ Ph.D. in Criminal Law and Criminology, Faculty of Law and Political Science, University of Tehran, Tehran, Iran.

² Assistant Prof, Department of Criminal Law and Criminology, Faculty of Law and Political Science, University of Tehran, Tehran, Iran.

Article Info	Abstract
Original Article	In the foreseeable future, with the ever-increasing advancements in emerging technologies, Artificial Intelligence (AI) will evolve from a weak to a strong level, reaching a stage of thinking, learning, and performance that allows it to be not only human-like but also autonomous and capable of acting in unpredictable ways. Such a capacity does not insulate this emerging agent, which can be termed an <i>Autonomous Artificial Intelligence Agent</i>, from the potential to cause harm to individuals, despite its reduced proneness to error. Consequently, the necessity of codifying legal rules concerning harms caused by these agents raises the challenge of their responsibility for committing crimes and inflicting damage. Therefore, this research adopts an analytical approach grounded in legal principles and developments regarding criminal and civil liability, to address three fundamental components of liability: legal personhood, conduct, and mental state. Ultimately, it concludes that with the possibility of granting legal personhood to the Autonomous Artificial Intelligence Agent, in addition to natural and legal persons, holding it liable for causing damage does not face significant obstacles. However, due to lingering doubts about the blameworthiness of these agents concerning criminal liability, adjudicating their responsibility for committing crimes remains challenging.
Received: 09-06-2024	
Accepted: 31-08-2024	
Keywords: Artificial Intelligence Legal Personhood Criminal Liability Civil Liability	
*Corresponding author e-mail: y.khademi@ut.ac.ir	
How to Cite: Khademi Shoormasti, Y., & Yekrangi, M. (2026). The Requirements of Liability of Artificial Intelligence in the Light of Criminal and Civil Law. <i>Modern Technologies Law</i> , 7(13), 21-41.	
Published by University of Science and Culture https://www.usc.ac.ir Online ISSN: 2783-3836	

1. Introduction

In the current world, individuals engage with Artificial Intelligence (AI) from the moment they begin their daily routines. This extensive integration of AI into human life has unfolded with such subtlety that individuals are neither willing nor able to disengage from it. In consequence, this technological phenomenon has garnered increasing attention across various disciplines most importantly within the humanities prompting a range of profound philosophical, ethical, social, psychological, and legal inquiries.

Despite the pervasive utilization of AI across multiple fields, a singular, universally accepted definition remains elusive. Nevertheless, a commonly cited conceptual thread among prevailing definitions one which necessitates the legal community's attention is AI's capacity for cognitive processing and behavior modeled on human thought and action. This shared feature underpins the inclusion of AI within the legal sphere as an inevitable legal subject. However, because AI exhibits this capacity to varying degrees, it encompasses a broad spectrum of systems, thereby complicating legal engagement with such entities. Each type of AI within this spectrum, from weak to strong, may require a distinct and specific legal response, in contrast to the legal system's foundational reliance on general and uniformly applicable norms.

This study does not limit AI to a mere tool for performing human tasks or inflicting harm; rather, it considers AI as an *Autonomous Artificial Intelligence Agent*, a form of intelligent entity that is capable of learning and perceiving from its environment, and of acting autonomously, independently of its developer, manufacturer, or user. Through such autonomous operation, an AI agent, much like a human actor, may exert influence upon its environment and cause harm to others. Although such agents have yet to achieve widespread deployment, it is anticipated that they will soon become integral to human social life. Their less advanced predecessors have already permeated human existence, and it is perhaps AI systems based on large language models (LLMs) that have attracted the most public attention. These distinctive characteristics of AI agents make it both necessary and challenging to identify legal responsibility attributable to them.

Consequently, the urgent question facing contemporary legal discourse is: what prerequisites must be satisfied for AI to enter the legal domain as a responsible agent? Moreover, can AI, within both civil and criminal frameworks, be deemed to possess the requisite legal capacity and competence to bear duties and liabilities? Notably, recognizing an intelligent agent as a subject of legal responsibility presupposes the necessity of legal legitimacy as a fundamental component of such responsibility. In the absence of this legitimacy, AI would remain an unrecognized and, indeed, an illegitimate entity in the eyes of the state, thereby shifting the burden of responsibility back onto its developers and providers. This issue obviously lies beyond the technical scope of the present study.

Although domestic studies contains scientific sources addressing AI responsibility—including translated works and those specifically examining civil liability, such as Rajabi's 2019 article, "*Liability of Artificial Intelligence: The Reflection of Developments in the Liability Rules*," and Takhshid's 2021 article, "*An Introductory Study on the Challenges of Artificial Intelligence in Tort Law*," as well as studies on criminal liability, including Khademi's 2019 master's thesis, "*Feasibility of Criminal Liability and Punishment of Artificial Intelligence*," and the 2024 article by Amirian Farsani and Hosseini, "*Challenges and Obstacles of Criminal Liability in Robots with Artificial Intelligence Capabilities*," a comprehensive comparative analysis of both civil and criminal liability, particularly about the conditions and prerequisites for recognizing legal responsibility, remains essential. In the foreign context, most scholarly discussions tend to focus exclusively on one of these two areas of liability, often about particular legal systems or in connection with moral responsibility, rather than engaging in a comparative, prerequisite-focused legal analysis.

2. Methodology

In this paper, data collection was conducted through a comprehensive review of academic sources, including books, scholarly articles, theses, and research reports obtained from reputable scientific and library-based

databases. Given the innovative, technological, and dynamic nature of the subject matter, the research also incorporated online resources to ensure familiarity with the latest developments and advancements in the field of artificial intelligence. This approach enabled the researchers to examine relevant legal data and doctrines in the context of foreign experiences, particularly regarding the formulation of laws and regulations governing AI.

Considering the subject at hand, and the necessity of addressing both the civil and criminal elements of liability, the research design is structured as follows: in the first section, the general prerequisite for the recognition of legal responsibility namely, *legal personhood* is analyzed as a foundational condition for assigning liability to any legal actor. The second section examines the feasibility of attributing the *actus reus* (material element) to artificial intelligence, while the third section evaluates the necessary conditions for establishing the *mens rea* (mental element) in relation to these emerging intelligent agents.

Throughout all discussions, civil and criminal liability are comparatively examined side by side, in order to identify points of convergence and divergence between the two domains. This comparative approach is adopted to forecast a coherent and legally sound framework for the inclusion of artificial intelligence within the legal order and for the imposition of liability upon such entities.

3. Results and Discussion

In contemporary times, given the expansion of research and the rapid advancements in developing systems that are purportedly becoming increasingly similar to human thinking and behavior in both method and mechanism, the necessity for close interaction between artificial intelligence and the legal system is felt more than ever. Among the various intersections between law and AI, the most striking is the consideration of artificial intelligence as a new subject of rights and duties within a legal framework.

AI agents, as certain level of advanced systems, are predicted to operate autonomously learning from and adapting to their environment in pursuit of their own objectives and interests. As such, they may cause harm to surrounding entities in the course of achieving these aims. This reality necessitates the identification of one or more responsible parties to bear liability for the damages caused, whether in the domain of criminal law or civil law.

By examining the foundations of civil and criminal liability, this study concludes that a fundamental prerequisite for the imposition of legal responsibility on AI agents independent of human agents—is the recognition of *legal personhood* for such entities. This legal personality entails a set of rights and entitlements that serve as the basis for imposing duties and obligations. Therefore, it is imperative to move beyond the traditional, human-centered view of legal personhood and to embrace novel perspectives that transcend the dichotomy of *personhood* versus *thinghood* in the legal realm.

Regarding the attribution of the material element of harm whether through act or omission to AI agents, the research finds that such attribution does not encounter the same doctrinal difficulties typically associated with legal persons, animals, or inanimate objects. First, an intelligent agent constitutes a tangible and non-fictional entity: while it may lack a robotic form or physical mobility, and may act solely within virtual or digital environments, it still operates through information and communication technologies and has a direct, perceptible impact on its surroundings an impact that can be observed through human senses. This differs fundamentally from legal persons, whose existence is entirely a matter of legal fiction and abstraction.

Moreover, the limitations traditionally raised about animals, objects, or even narrow AI systems in attributing actions to the entity itself do not apply in the same way to AI agents. In those earlier cases, the law could not recognize such actors as legal subjects in their own right and instead had to rely on an associated human actor to establish liability. However, intelligent agents with their physical and perceivable existence, and with legal personality, rights, and duties recognized alongside humans can be treated as independent legal subjects. As such, the consequences of their actions may be pursued independently, without reliance on other legal actors.

Finally, with regard to the *mental element*, which arguably presents the greatest legal challenge in the context of AI, civil liability appears less problematic. This is due to the shift from moral to social standards of fault, and

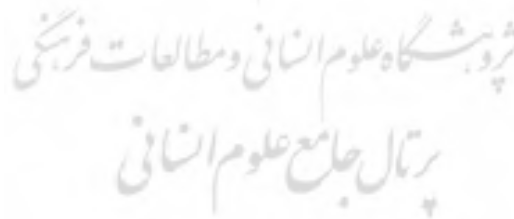
even the elimination of fault as a requirement in certain categories of civil liability. As a result, the ethical question of whether AI agents can be morally blameworthy is less pressing in this domain, making the recognition of civil liability more straightforward than that of criminal liability.

Conversely, criminal liability because of its inherent link with moral responsibility may encounter greater resistance among legal scholars when it comes to assigning such responsibility to AI agents. Nevertheless, abandoning a purely traditional perspective and embracing more contemporary theories can pave the way forward. In particular, the denial of awareness, intent, or volition in AI agents becomes increasingly untenable in light of the development of mechanisms such as *transparency* and *explainability* in artificial intelligence systems.

4. Conclusions and Future Research

Overall, it appears that what generates skepticism regarding the personal and direct liability of AI agents is neither the inherent possibility of attributing the material element of an offense nor even their capacity for blameworthiness. Rather, it is the legal system's perspective on these emerging actors as potential legal persons that remains the core issue. Therefore, a legal system should formally recognize intelligent agents as legal persons endowed with a defined and proportionate set of rights, duties, and obligations then such agents may indeed be held accountable for their actions and behaviors, including the consequences that follow. Nonetheless, the recognition of legal personhood is not merely a legal issue; it intersects deeply with philosophical, social, and ethical debates surrounding intelligent agents.

Accordingly, despite the current legal literature's predominant focus on the responsibility of human actors rather than that of intelligent agents, it must not be overlooked that if the legal system merely observes societal changes without actively adapting to them, it will eventually lose its effectiveness in governance and in maintaining order. In such a case, the law's inability to respond to rapidly emerging social phenomena will not only diminish its practical utility, but may also impose substantial and at times irreparable costs on both the legal system and society at large.





حقوق فناوری‌های نوین

<http://doi.org/10.22133/mtlj.2025.461599.1334>

مقتضیات مسئولیت عامل دارای هوش مصنوعی در پرتو حقوق کیفری و مدنی

یاسمن خادمی شورمستی^{۱*} (ID)، محمد یکرنگی^۲ (ID)

^۱ دکتری حقوق جزا و جرم‌شناسی، دانشکده حقوق و علوم سیاسی، دانشگاه تهران، تهران، ایران.

^۲ استادیار گروه حقوق جزا و جرم‌شناسی، دانشکده حقوق و علوم سیاسی، دانشگاه تهران، تهران، ایران.

اطلاعات مقاله	چکیده
مقاله پژوهشی	با توجه به پیشرفت‌های روزافزون در حوزه فناوری‌های نوین، در آینده نه‌چندان دور، هوش مصنوعی با عبور از سطح ضعیف به قوی، به مرتبه‌ای از تفکر، یادگیری و عملکرد خواهد رسید که می‌تواند نه صرفاً شبیه به انسان، بلکه هم‌پای او خودمختار و به‌نحوی پیش‌بینی‌ناپذیر عمل کند. چنین ظرفیتی این کنشگر نوظهور را، که از آن می‌توان تحت‌عنوان «عامل دارای هوش مصنوعی» یاد کرد، از ایراد ضرر به اشخاص به‌رغم خطاپذیری کمتر، مصون نمی‌سازد. بدین ترتیب، ضرورت تنظیم قواعد حقوقی پیرامون آسیب‌رسانی از سوی این عامل‌ها، چالش مسئولیت‌پذیری آن‌ها در قبال ارتکاب جرم و ایراد خسارت را پیش روی ما قرار می‌دهد. ازاین‌رو، پژوهش حاضر در پاسخ به مسئله فوق، با اتخاذ رویکرد تحلیلی و بر پایه مبانی و اصول حقوقی و تحولات ایجادشده درخصوص مسئولیت کیفری و مدنی، ضمن پرداختن به سه مؤلفه اساسی در مسئولیت، یعنی شخصیت قانونی، رفتار و وضعیت روانی، درنهایت به این نتیجه می‌رسد که با امکان‌پذیربودن اعطای شخصیت قانونی به عامل دارای هوش مصنوعی در کنار اشخاص حقیقی و حقوقی، مسئول‌شناختن آن به‌موجب ایراد خسارت با مانع جدی مواجه نبوده؛ اما درخصوص مسئولیت کیفری، به‌دلیل تردیدها در سرزنش‌پذیری این عامل‌ها، حکم به مسئولیت آن‌ها به‌موجب ارتکاب جرم همچنان با دشواری روبه‌روست.
تاریخ دریافت: ۱۴۰۳/۰۳/۲۰	
تاریخ پذیرش: ۱۴۰۳/۰۶/۱۰	
واژگان کلیدی: هوش مصنوعی شخصیت قانونی مسئولیت کیفری مسئولیت مدنی	
*نویسنده مسئول رایانامه: y.khademi@ut.ac.ir	

نحوه استناددهی:

خادمی شورمستی، یاسمن، و یکرنگی، محمد (۱۴۰۵). مقتضیات مسئولیت عامل دارای هوش مصنوعی در پرتو حقوق کیفری و مدنی. حقوق فناوری‌های نوین، ۷(۱۳)، ۲۱-۴۱.

ناشر: دانشگاه علم و فرهنگ <https://www.usc.ac.ir>

شاپای الکترونیکی: ۲۷۸۳-۳۸۳۶

مقدمه

در دنیای کنونی، هر شخص از زمانی که روز خود را آغاز می‌کند، در تعامل با هوش مصنوعی^۱ قرار می‌گیرد؛ از مدیریت فعالیت‌های روزمره با استفاده از ابزارهای هوشمند شخصی و خانگی گرفته تا مرور اخبار در رسانه‌ها و انجام معاملات سودده در بازارهای مالی دیجیتال. این تسخیر حیات با هوش مصنوعی آنچنان آرام و دل‌نشین صورت می‌پذیرد که انسان نه می‌خواهد و نه قادر است از آن روی‌گردان شود. همین امر موجب جلب توجه حوزه‌های مختلف دانش، به‌ویژه علوم انسانی به این فناوری نوین و پدیدآمدن پرسش‌های تأمل‌برانگیز فلسفی، اخلاقی، اجتماعی، روان‌شناختی و حقوقی پیرامون آن شده است.

هوش مصنوعی از جمله مفاهیمی است که با وجود کثرت استعمال در دانش‌های گوناگون، تعریف واحد و مشخصی برای آن ارائه نشده و شاید همین دشواری در تعریف است که بر ظهور چالش‌ها پیرامون آن دامن زده است. این دانش‌واژه علاوه بر آن که یک رشته مطالعاتی علمی معرفی می‌شود که در این زمینه نیز تعریف‌های بسیار و تاحدی متفاوت برای آن بیان شده است (Russell & Norvig, 2011, p. 18)، بر سیستم‌ها و موجودیت‌هایی^۲ نیز، که حاصل پژوهش در این رشته مطالعاتی هستند، اطلاق می‌شود. فارغ از اختلاف در تعاریف هوش مصنوعی، خواه در قالب رشته مطالعاتی و خواه در قالب موجودیت، آنچه بنابه دیدگاه مشهور، فصل مشترک همه آن تعاریف عنوان می‌شود^۳ و همین ویژگی است که درواقع ورود موجودیت‌های دارای هوش مصنوعی را به عرصه حقوق به‌منزله مخاطب ناگزیر می‌سازد؛^۴ قدرت بر تفکر و رفتار بر محوریت تقلید و پیروی از تفکر و رفتار انسانی است؛ اعم از گرایش‌ات مبتنی بر تفکر و کنشگری عقلانی و منطقی و گرایش‌ات مبتنی بر تفکر و کنشگری شبه‌انسانی (نه لزوماً منطقی) (Taghani, 2017, p. 28)؛ بنابراین می‌توان گفت منظور از هوش مصنوعی به‌منزله موجودیتی جدید، سیستمی است که رفتارهایی مشابه رفتار انسان انجام می‌دهد (Čerka et al, 2015, p. 378)، به اهدافی پیچیده دست می‌یابد (Tegmark, 2018, p. 72) و با حس کردن محیط اطراف و پردازش آن به‌نحو مناسب، در محیطی که در آن قرار دارد کنشگری می‌کند (Stone et al, 2016, p. 12-13).

باین‌حال، باید توجه داشت که برخورداری از چنین توانمندی‌هایی از آنجاکه بر طیفی از درجات گوناگون قرار می‌گیرد، هوش مصنوعی را به انواع مختلفی تقسیم می‌کند و در نتیجه، مواجهه حقوق با آن را با مشکل همراه می‌سازد؛ چراکه هرگونه از این سیستم‌ها در این طیف، برخورد متفاوتی را می‌طلبد و این درحالی است که اساس حقوق بر قاعده‌گذاری کلی است.

سیستم‌های دارای هوش مصنوعی، از جهات مختلف فناورانه مانند نوع الگوریتم‌های یادگیری، فرایند پیشرفت و توانایی‌های شناختی، به گونه‌های متفاوتی تقسیم می‌شوند که اگرچه تمامی آن‌ها در تلاش برای رفتار و به‌کارگیری اعمال شناختی مشابه انسان مشترک هستند، اما تفاوت‌های چشم‌گیری با یکدیگر دارند^۵. مهم‌ترین تقسیم‌بندی در این نگاه بر پایه میزان توانمندی هوش مصنوعی در یادگیری، «هوش مصنوعی قوی» را، که مقصود اصلی پژوهش حاضر است، در برابر «هوش مصنوعی ضعیف» قرار می‌دهد. این نوع از هوش مصنوعی نه صرف یک ابزار، بلکه سیستمی است که قادر است ذهن تلقی شود؛ در عین اینکه توانایی درک و سایر حالات شناختی را دارد. به عبارت دیگر، هوش مصنوعی

1: Artificial intelligence (AI)

2: Entity

۳ در دیدگاه اقلیت، هوش مصنوعی صرفاً تلاش برای ساخت سیستم‌هایی است که می‌توانند مسائلی کاربردی را برای انسان حل کنند؛ بدون آن‌که در رویای ساخت سیستم‌هایی همانند انسان به سر ببرند (رهمنون، ۱۳۸۴، ص ۱۵).

۴ مقصود در این پژوهش، مطالعه پیرامون عامل‌های دارای هوش مصنوعی است؛ از آن جهت که می‌توانند مانند انسان موضوع مقررات و قواعد حقوقی باشند، اما ارتباط حقوق با رشته هوش مصنوعی به‌طور عام در قالب‌های دیگر، مانند استفاده از ابزارهای هوشمند در ساختار تقنینی، حقوقی و قضایی نیز شکل می‌گیرند. در این زمینه ر.ک: ابوذری، مهرنوش (۱۴۰۱)، حقوق و هوش مصنوعی، ج ۲، تهران: نشر میزان.

۵ هوش مصنوعی بنابه تقسیم‌بندی‌های مشهور به هوش مصنوعی قوی (Strong AI) و ضعیف (Weak AI) از یک‌سو، ماشین‌های واکنشی (Reactive Machines)، حافظه محدود (Limited Memory)، نظریه ذهن (Theory of Mind) و ماشین‌های خودآگاه (Self-awareness) از یک سو و همچنین، هوش مصنوعی محدود (AND)، هوش عمومی مصنوعی (AGI) - که کاملاً مشابه هوش انسانی است - و هوش فراتر از انسانی (ASI) از سوی دیگر تقسیم می‌شوند که البته هم‌پوشانی‌هایی نیز با یکدیگر دارند.

ضعیف دارای الگوریتم‌هایی است که رفتار سیستم را از طریق برنامه‌های ازپیش‌تعیین‌شده کنترل می‌کند و نتایج عملکرد آن تا حد بسیار زیادی پیش‌بینی‌پذیر است؛ درحالی‌که درمقابل، هوش مصنوعی قوی، الگوریتم‌هایی را برای یادگیری ماشینی^۱ داراست که سیستم را قادر می‌سازد با یادگیری از تجربیات گذشته، به اصلاح عملکرد و بازبینی در الگوریتم‌های خود اقدام کند (Searle, 1980, p. 417). درواقع، هوش مصنوعی قوی از سطحی از خودمختاری^۲ بهره‌مند است؛ نه‌تنها به این معنا که تصمیم‌گیری‌هایی مستقل از مداخله و کنترل انسان انجام می‌دهد، بلکه در معنای بالاتر، به این مفهوم که تصمیماتش برای انسان به‌طور کامل پیش‌بینی‌پذیر نیست (Tasioulas, 2019, p. 52).

با توجه به آنچه گفته شد، در این پژوهش منظور از هوش مصنوعی، صرفاً ابزاری برای انجام فعالیت‌های انسانی و ایجاد آسیب نیست^۳؛ بلکه هوش مصنوعی به‌عنوان یک «عامل»^۴ است که می‌تواند با درک محیط پیرامون و یادگیری از آن به‌نحوی مستقل از ارائه‌دهنده، سازنده و کاربر و به‌شکلی خودمختار در محیط پیرامونش اثرگذار باشد و به‌موجب این کنشگری، همان‌گونه که یک انسان می‌تواند آسیب‌هایی را به دیگران وارد آورد، موجب بروز صدمات به افراد شود.

لذا زین پس، از این نوع از موجودیت‌های هوشمند تحت‌عنوان «عامل دارای هوش مصنوعی یا هوشمند» یاد خواهیم کرد؛ عامل‌هایی که هرچند هنوز در این سطح فراگیر نشده‌اند، اما پیش‌بینی می‌شود که در آینده نزدیک می‌توانند به جزئی جدایی‌ناپذیر از زندگی اجتماعی انسان بدل شوند؛ همچنان‌که هم‌نوعان کمترپیشرفته آنان امروزه زندگی انسانی را به انقیاد خود درآورده‌اند و شاید بیش از همه، هوش مصنوعی مبتنی بر مدل‌های زبانی^۵ است که توجه عموم را به خود جلب کرده است.

اما در کنار دسته‌بندی‌های فناورانه فوق و تمرکز نوشتار حاضر بر هوش مصنوعی قوی، ضروری است طبقه‌بندی دیگری را نیز، این بار از منظر کارکرد و میزان خطرآفرینی این عامل‌های هوشمند در جامعه مدنظر قرار داد که چندسالی است موردتوجه سیاست‌گذاران واقع شده است.^۶ برای مثال اتحادیه اروپا، احکام اخیر خود^۷ را به‌لحاظ نظری، بر طبقه‌بندی چهارگانه هوش مصنوعی (۱) دارای خطر غیرقابل‌پذیرش؛ (۲) پرخطر؛ (۳) دارای خطر خاص؛ و (۴) کم‌خطر پایه‌گذاری کرده^۸ که به‌ترتیب آن‌ها را ممنوع، مشروط به کنترل، مشروط به شفافیت و غیرمشروط دانسته است.^۹ همچنین برزیل در لوایح مرتبط با هوش مصنوعی،^{۱۰} بر ممنوعیت به‌کارگیری «هوش مصنوعی دارای خطر مفرط»^{۱۱} و ازسوی دیگر، کانادا در لایحه «قانون داده و هوش مصنوعی»^{۱۲}، بر تنظیمگری «هوش مصنوعی تأثیرگذار»^{۱۳} تأکید داشته‌اند. چنین طبقه‌بندی‌هایی ازآنجایی برای ما اهمیت می‌یابد که سخن از عامل هوشمند به‌عنوان شخصیتی مسئول، لزوم مشروعیت حقوقی را نیز به‌منزله مؤلفه ضروری مسئولیت

1: Machine Learning

2: Autonomy

۳. ر.ک: رجبی، عبدالله (۱۳۹۸)، ضمانت در هوش مصنوعی، مطالعات حقوق تطبیقی، دوره ۱۰، شماره ۲، صص. ۴۴۹-۴۶۶.

4: Agent

۵. سیستم‌های مبتنی بر مدل‌های زبانی بزرگ (Large Language Model) سیستم‌هایی همچون ChatGPT و Google Gemini هستند که با بهره‌گیری از یادگیری ماشینی نظارت‌شده (Supervised Learning) و تقویتی (Reinforcement Learning) به‌عنوان یک بات گفت‌وگو استفاده می‌شوند.

۶. در «سند ملی هوش مصنوعی جمهوری اسلامی ایران» مصوب ۱۴۰۳/۰۳/۲۹ شورای عالی انقلاب فرهنگی نیز بر لزوم ایجاد زیرساخت‌های حقوقی برای مواجهه با مخاطرات هوش مصنوعی ازجمله «تعیین مسئولیت‌پذیری فردی و اجتماعی هوش مصنوعی خودمختار» تأکید شده است.

۷. نخست، قانون هوش مصنوعی (Artificial Intelligence Act) (European Parliament, 2024, 13 Jun.) است که به تصویب رسیده و دوم، دستورالعمل مسئولیت هوش مصنوعی (Artificial Intelligence Liability Directive) (European Commission, 2022, 28 Sept.) است که تاکنون سرانجام نیافته است.

8: 1. Unacceptable risk; 2. High risk; 3. Specific risk; and 4. Minimal risk.

9. “Artificial Intelligence—Questions and Answers.” European Commission, 1 Aug. 2024. Available at: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683 Accessed 3 Aug. 2024.

10. “Access Alert | Brazil’s New AI Bill: A Comprehensive Framework for Ethical and Responsible Use of AI Systems.” Access Partnership, 5 May 2023. Available at: <https://accesspartnership.com/access-alert-brazils-new-ai-bill-a-comprehensive-framework-for-ethical-and-responsible-use-of-ai-systems/> Accessed 10 Apr. 2024.

11: Excessive risk

12. “The Artificial Intelligence and Data Act (AIDA) – Companion Document.” 13 Mar. 2023. Available at: <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document/> Accessed 10 Apr. 2024.

13: High-impact AI

شخصی این کنشگر با خود به همراه می‌آورد؛ زیرا در غیر این صورت هوش مصنوعی صرفاً موجودیتی به رسمیت شناخته نشده و بالاتر از آن، غیر مشروع از سمت دولت خواهد بود که ما را دوباره به سمت مسئولیت ارائه‌دهندگان و سازندگان آن سوق می‌دهد؛ مسئله‌ای که از نظر تخصصی از موضوع پژوهش پیش رو خارج است.

در نهایت، شایان ذکر است که بحث از مسئولیت عامل دارای هوش مصنوعی، فارغ از مسئله ضمانت اجرای کیفری و مدنی، قابل اعمال بر این عامل‌هاست. از این رو، پرسش از انواع، چگونگی و شیوه اعمال ضمانت اجراها بر عامل هوشمند یا امکان‌سنجی بهره‌گیری از نهادهای نیابتی، در عهده‌گیری این ضمانت اجراها از سوی اشخاص دیگری غیر از عامل دارای هوش مصنوعی، هرچند موضوع بسیار حائز اهمیت است، اما آشکارا خارج از قلمرو موضوعی این پژوهش قرار می‌گیرد. بدین ترتیب، ضمن گزینش عامل هوشمند - در قالب هوش مصنوعی قوی که دولت واجد مشروعیت لازم برای کنشگری در جامعه می‌شناسد - به منزله موضوع تحقیق حاضر، ابتدا امکان بهره‌مندی این عامل‌ها از شخصیت قانونی مانند سایر اشخاص حقیقی و حقوقی بررسی می‌شود (بند اول) و سپس قابلیت انتساب رفتار و نتایج آن به عامل دارای هوش مصنوعی (بند دوم) و در نهایت، اسناد عنصر روانی لازم برای مسئولیت مورد مذاقه قرار می‌گیرد (بند سوم).

۱. بند اول: شخصیت قانونی^۱ عامل دارای هوش مصنوعی

مهم‌ترین مسئله در موضوعات حقوقی مرتبط با هوش مصنوعی این پرسش است که عامل هوشمند در یک نظام حقوقی چه جایگاهی دارد. در مقام شخصیتی مستقل، یک عامل و کنشگر دارای صلاحیت برای برخورداری از حقوق و تکالیف قانونی (Pagallo, 2013, p. 37) و حتی حقوق شهروندی مشابه اشخاص حقیقی (Pagallo, 2013, p. 54) یا صرفاً ابزار و منبعی برای مسئولیت و تعهد سایر کنشگران مرتبط با آن، مانند ارائه‌دهندگان، سازندگان و کاربران. به طور سنتی، یک شخص قانونی به موجب حقوقی که از یک نظام حقوقی دریافت می‌کند، دارای این اهلیت و صلاحیت خواهد بود که طرح دعوا کند، به نام و به حساب خود اموالی را دارا شود، طرف قراردادهای قرار گیرد و همچنین، از مصونیت‌ها و حمایت‌های گوناگونی برخوردار شود؛ مانند حق بر آزادی (Chopra & White, 2011, p. 154)؛ هرچند تمامی اشخاص قانونی، الزاماً حقوق و تکالیف مشابه و یکسانی را نخواهند داشت. به عقیده برخی، حتی اگر صرفاً در حقوق شکلی مربوط به دعاوی، با عامل دارای هوش مصنوعی به عنوان شخصیتی غیر از یک مال یا ابزار برخورد شود، پیش از آن‌که به طور رسمی دارای شخصیت قانونی شود، این نوع برخورد ناگزیر، موجب تحول حقوق ماهوی این عامل‌ها برای برخورداری از حقوق بنیادی خواهد شد؛ چنان‌که در مورد اشخاص حقوقی این اتفاق رخ داد (Michalski, 2018, p. 1070).

پیش از امکان‌سنجی ظرفیت عامل هوشمند در داشتن حق و تکلیف، شناسایی شخصیت قانونی برای این عامل به عنوان پیش شرط مسئولیت کیفری و مدنی ضروری و نیازمند تبیین است. به طور کلی، در یک نظام حقوقی، فقط در صورتی می‌توان برای یک شخص تکالیفی مقرر داشت و وی را بابت اعمالی که انجام می‌دهد مسئول دانست که پیش‌تر شخصیت مستقل وی در آن نظام شناسایی شده باشد (Chakrabarty, 2023, p. 4) & Baral. به بیان دیگر، شناسایی شخصیت قانونی و هویت مستقل برای عامل‌های دارای هوش مصنوعی و به تبع آن قائل شدن حقوق و تکالیف برای آن‌ها از منظر حقوقی، به این دلیل اهمیت می‌یابد که مسئول شناختن عامل آسیب‌رسان و خاطی، همواره متضمن محدودیت در حقوق مادی و حتی معنوی و تحمیل تکلیف و الزامی بر عهده اوست که بدون احراز پیشینی یک سری حقوق و امکان تحمل تکالیف و تعهدات، تحمیل مسئولیت و اعمال ممنوعیت و محدودیت به لحاظ منطقی امری ناممکن خواهد بود؛ زیرا تا حقی وجود نداشته باشد، اعمال محدودیت و ضمانت اجرا نسبت به آن نیز سالبه به انتفاء موضوع خواهد بود؛ لذا این‌که به دور از شناسایی شخصیت قانونی، به دنبال مسئولیت‌پذیری عامل هوشمند از طریق قیاس آن با برده، صغیر یا مجنون باشیم (Hekmatnia et al., 2019, pp. 253-256)، به نظر شیوه‌ای راهگشا نخواهد بود؛ زیرا موارد اخیر نیز به تمامی شخصیت داشته و تنها کیفیات و حدود حقوق و تکالیفی که به موجب شخصیت قانونی خود

1: Legal Personhood

از آن بهره‌مند می‌شوند متفاوت است؛ نه آن‌که احتساب عامل دارای هوش مصنوعی به‌عنوان برده، صغیر یا مجنون، راهبردی هم‌عرض با احتساب عامل هوشمند به‌عنوان یک شخص قانونی باشد. بنابراین، نباید از یاد برد که اعطای شخصیت قانونی به موجودیت‌های فعال از منظر حقوقی در جامعه، نه از باب امتنان دولت‌ها، بلکه یک ضرورت حقوقی است.

اکنون با دریافت این ضرورت، باید دید عامل دارای هوش مصنوعی چگونه در راستای شخصیت‌پذیری در یک نظام حقوقی گام برمی‌دارد. تبیین این امر نیازمند کنارهم قراردادن نظریات سنتی و نو در باب معیار و آثار شخصیت قانونی و سنجش ظرفیت عامل هوشمند در انطباق با آن‌هاست. در دیدگاه سنتی، شخصیت قانونی با انسان بودن به‌عنوان موجودیت دارای هویت مستقل و برخوردار از اهلیت برای به‌دست‌آوردن حق و تکلیف پیوند خورده است. در این رویکرد، هر موجودیت دیگری که بخواهد شخصیت قانونی به دست آورد باید در سنجش با انسان قرار گرفته و در صورتی که بتواند به چنین اهلیتی متصل شود، شناسایی شخصیت برای وی بلامانع خواهد بود (Dyschkant, 2015, pp. 2077-2078)؛ چنان‌که در مورد اشخاص حقوقی نیز ضمن پابندی به این امر و با قدری انعطاف در تحلیل، به‌دلیل اعتبار آن‌ها به‌عنوان شرکت جمعی از انسان‌ها و دارایی‌هایشان، شخصیت‌پذیری ممکن شد. مطابق این دیدگاه، اطلاق شخصیت قانونی به موجود غیرانسانی همچون عامل هوشمند، غیرواقعی جلوه خواهد کرد؛ مگر آن‌که چنین موجودیتی دارای خصایص انسانی و به‌ویژه هوش و اراده باشد (Solum, 1992, p. 1239). این درحالی است که گروهی معتقدند هوش مصنوعی هیچ‌گاه نخواهد توانست دارای روح، آگاهی، قصدمندی، احساسات و اراده آزاد شود و بنابراین، نمی‌تواند شخصیت قانونی داشته باشد (Monterossi, 2020, p. 3). ادعایی که دست‌کم در مورد قصدمندی، آگاهی و اراده آزاد، حتی در خصوص هوش مصنوعی فعلی نیز چندان با واقع منطبق نیست. چنان‌که در باب اقسام آن گذشت، به نظر می‌رسد عامل هوشمند قادر است با کسب هوشی معادل و به نظر برخی حتی فراتر از انسان و اراده آزاد و خودمختار در آینده عمل کند. این نکته نیز باید مدنظر قرار گیرد که قائل شدن حقوق قانونی برای موجودیت‌های دارای شخصیت قانونی مانند حق بر آزادی بیان، حمایت یکسان در برابر قانون و مانند آن، مستلزم قائل شدن حقوق اخلاقی برای این اشخاص نیست؛ هرچند ممکن است در مواردی هم‌پوشانی وجود داشته باشد (Solum, 1992, p. 1240). درمقابل این نگرش، برخی از نظریات جدید در واقع تلقی سنتی و به‌ویژه انسان‌محور درباره شخصیت قانونی و لزوم ابتدای آن بر اهلیت برخوردار از حق و تکلیف را یگانه برداشت از مفهوم شخصیت قانونی نمی‌دانند. بنابر دیدگاه نو، شخصیت قانونی به معنای برخورداری از برخی یا تمام استحقاقات قانونی است که به برخی یا همه، به‌دلیل موقعیتشان به‌منزله یک شخص قانونی اعطا می‌شود. گروهی ممکن است از تمامی این استحقاقات برخوردار باشند و عده‌ای دیگر از برخی؛ اما این به معنای فقدان شخصیت قانونی دسته‌آخر نیست (Kurki & Pietrzykowski, 2017, p. 83). برپایه این نظر، شخصیت قانونی به نفس اهلیت حق و تکلیف بر نمی‌گردد؛ بلکه به دسته‌ای از امتیازات فعال و منفعلی برگشت دارد که شخص ممکن است از برخی یا همه بهره‌مند شود و در تمامی این حالات شخص قانونی محسوب شود؛ اما یکی با استحقاقات و بار مسئولیت بیشتر و یکی کمتر (Kurki & Pietrzykowski, 2017, p. 84) و به همین دلیل، مرز قاطعی میان شخصیت و عدم شخصیت نمی‌توان ترسیم کرد. در همین راستا، برخی به این سمت گام برداشته‌اند که ترسیم دوگانه شخص انگاری و شیء‌انگاری^۱، که بر اومانیسیم حقوقی دامن می‌زند، در آینده پاسخ‌گوی نیازهای ما در جامعه نخواهد بود (Kurki & Pietrzykowski, 2017, p. 65). این گروه برخلاف نگرش سنتی، دیدگاهی کارکردی به شخصیت قانونی داشته و از باب توزیع مناسب‌تر مسئولیت، بر اعطای شخصیت قانونی به عامل‌های دارای هوش مصنوعی تأکید دارند (Novelli, 2023, p. 1347)؛ بنابراین قائلان دیدگاه اخیر بر این باورند که حتی اگر براساس گرایش سنتی نتوان عامل هوشمند را شخصی قانونی محسوب کرد، هوش مصنوعی براساس شناسایی یک‌سری امتیازات، ولو کاملاً به حد اشخاص قانونی حقیقی نرسد، می‌تواند از شخصیت قانونی برخوردار شود (Asaro, 2007, p. 22).

فارغ از نگرش‌های سنتی و نو، واقعیت انکارناپذیر آن است که شناسایی شخصیت قانونی، علاوه بر مقتضیات فنی و حقوقی، نیازمند بستر اجتماعی لازم است. درواقع اعطای شخصیت قانونی ازسوی یک نظام حقوقی در پاسخ به فشارهای سیاسی، اخلاقی و حقوقی و برای کارکرد

1: Personhood and Thinghood

بهرتر نظام حقوقی در راستای سازگاری با واقعیات اجتماعی صورت می‌گیرد (Chopra & White, 2011, p. 155). واقعیتی که با درک اجتماعی ما از عامل دارای هوش مصنوعی به مثابه یک کنشگر ارتباط مستقیم دارد (Michalski, 2018, p. 1069). بنابراین، هنگامی لزوم واکنش مناسب نظام حقوقی پدیدار می‌شود که عامل هوشمند به عنوان یک سیستم اجتماعی فنی^۱ (Kudina & van de Poel, 2024, p. 2) با ورود به جامعه و زندگی اجتماعی انسان، به عنوان یک کنشگر، شخصیت و هویت مستقل از نگاه اعضای اجتماع درک شود.

در نهایت، امکان برخورداری هوش مصنوعی از شخصیت قانونی نه تنها به سادگی ردشدنی نیست؛ بلکه پیش‌تر جلوه‌هایی نیز در برخی کشورها داشته است. از جمله اولین و مهم‌ترین سند در ارتباط با اعطای یک شخصیت رسمی به عامل دارای هوش مصنوعی، مقررۀ ۲۰۱۷ پارلمان اروپا با عنوان «مقررات حقوق مدنی در خصوص ربات‌ها» است که از اعطای شخصیت الکترونیک به هوش مصنوعی در آینده نزدیک سخن می‌گفت (European Parliament, 2017, 16 Feb.). تا در صورتی که بتواند تصمیم‌هایی را به نحو خودمختار بگیرد و با دیگران مستقلاً تعامل داشته باشد، به موجب شخصیت الکترونیک، مسئول نتایج اعمال خود باشد که دوباره در سال ۲۰۲۰ نیز با شرح و تفصیل بیشتر به آن پرداخته شد (European Parliament, 2020, 20 Oct.). از سوی دیگر، در سال ۲۰۱۷ ربات انسان‌نمای پیشرفته‌ای به نام سوفیا، که قادر است به عنوان رباتی اجتماعی و با پردازش محیط با افراد به گفت‌وگو بپردازد، به عنوان نخستین ربات شهروند، تابعیت عربستان را کسب کرد^۲ و اندکی پس از آن، ژاپن نیز به شخصیت مجازی هوشمندی به نام میرای، که هفت سال داشت، اقامت خود را اعطا کرد^۳. این در حالی بود که این دو سیستم دارای هوش مصنوعی در قالب شهروندان هوشمند، بدون آن که مشخص شود چه حقوق و تکالیفی خواهند داشت، به یکباره به نظام حقوقی و شهروندی ژاپن و عربستان وارد شدند (Atabekov & Yastrebov, 2018, pp. 776-777).

بدین ترتیب، با امکان بهره‌مندی عامل دارای هوش مصنوعی از صلاحیت لازم برای داشتن حق و تکلیف، به ویژه با اتکا به نظریات نو-ولو آن که ممکن است از منظر مسائلی همچون مراتب و حدود شخصیت قانونی، اعتبار آن در قلمروهای حقوقی و آغاز و خاتمه آن، هم‌پایه و هم‌مرتبه با اشخاص حقیقی و حقوقی قرار نگرفته و احکام متفاوتی را ایجاب کند - ادامه پژوهش به امکان‌سنجی ارکان مسئولیت در خصوص این کنشگران نوپدید به عنوان عامل‌های خودمختار معطوف می‌شود.

۲. بند دوم: اسناد رکن مادی به عامل دارای هوش مصنوعی

پس از شخصیت قانونی، اولین قدم در شناسایی مسئولیت حقوقی عامل‌های هوشمند، امکان اسناد رکن مادی رفتار آسیب‌رسان به آن‌هاست؛ اعم از آن که به دنبال مسئولیت کیفری ناشی از زیان مجرمانه یا مسئولیت مدنی ناشی از زیان غیرقراردادی باشیم. انتساب رکن مادی در مسئولیت کیفری بدین معناست که رفتار مجرمانه از سوی مرتکب واقع شود؛ به نحوی که رابطه سببیت میان عمل مرتکب، اعم از فعل و ترک فعل و جرم مدنظر برقرار شود. مشابه همین امر در مسئولیت مدنی نیز وجود دارد؛ به صورتی که عمل زیان‌بار از سوی عامل زیان انجام گیرد، ضرری ناشی از عمل او به دیگری وارد آید و بتوان رابطه سببیت بین رفتار زیان‌بار و ضرر واردشده را احراز کرد. باید توجه داشت که عمل زیان‌بار در نتیجه تحولاتی که نظریات مسئولیت مدنی از نظریات «مبتنی بر تقصیر» تا نظریه «تضمین حق» پشت‌سر گذاشته است، الزاماً منوط به وجود تقصیر در مفهوم اخلاقی آن نیست؛ بلکه کافی است حقوق یک کشور، اعم از قانون، اخلاق و نظم عمومی، کاری را نامشروع بداند تا بگوییم رفتاری که انجام شده، رفتاری زیان‌بار بوده است (Katouzian, 2003, p. 311). در مقابل، رفتار آسیب‌رسان در مسئولیت کیفری اصولاً رفتار مادی صرف

1: Sociotechnical System (STS)

2: Weller, Chris. "Meet the First-ever Robot Citizen — a Humanoid Named Sophia That Once Said It Would 'destroy Humans.'" Business Insider, 27 Oct. 2017. Available at: www.businessinsider.com/meet-the-first-robot-citizen-sophia-animatronic-humanoid-2017-10 Accessed 10 Apr. 2024.

3: Cuthbertson, Anthony. "Tokyo: Artificial Intelligence 'Boy' Shibuya Mirai Becomes World's First AI Bot to Be Granted Residency." Newsweek, 6 Nov. 2017. Available at: www.newsweek.com/tokyo-residency-artificial-intelligence-boy-shibuya-mirai-702382 Accessed 10 Apr. 2024.

نیست و خطای کیفری عموماً ناظر به رفتاری است که ضرر یا امری غیراخلاقی را در بطن خود دارد و به سادگی از مفهوم تقصیر اخلاقی منفک نمی شود.

پیش از ورود به ابعاد حقوقی رکن مادی، پاسخ به یک پرسش بنیادی ضروری می نماید و آن این است که هوش مصنوعی با وجود داشتن توانمندی های شناختی مشابه و حتی بعضاً بهتر از انسان، چگونه ممکن است از منظر فنی عملکردی آسیب رسان داشته باشد؟ پاسخ به این مسئله از دو منظر درخور توجه است. نخست آن که هر روز بر تنوع و تکثر فعالیت هایی که هوش مصنوعی در آن به نقش آفرینی مشغول می شود افزوده می گردد؛ از فعالیت های شاید به ظاهر ساده تری همچون امور آموزشی، خدماتی و امدادی گرفته تا جراحی های پزشکی، امور نظامی و پلیسی، کمک به مقامات قضایی و پیشرانی در حکمرانی هوشمند.^۱ چنین فعالیت های گسترده ای که عمدتاً هم در تعامل مستقیم با انسان ها صورت می گیرد، بستر بالقوه ای برای وقوع زیان های مختلف مادی و معنوی را فراهم می آورد؛ به نحوی که حتی عامل هوشمند نیز با وجود خطاپذیری کمتر نسبت به انسان و اتفاقاً با توجه به عملکرد خودمختارانه، ساختار سخت افزاری و نرم افزاری، اشخاصی که با آن ها در تعامل است و محیط حضور و نوع فعالیتی که انجام می دهد از آن ها مصون نیست (Bertolini et al., 2016, p. 386).

دوم آن که با کاهش عملکرد خودکار^۲ و افزایش فعالیت خودمختار^۳ در انواع هوش مصنوعی، وجود همان بسترهایی که این سیستم ها به موجب آن ها همگان را در حیرت فرومی برند، انجام فعل یا ترک فعل مجرمانه و زیان بار را از سوی آن ها محتمل می گرداند؛ از جمله خطاهای ناشی از عملکرد برنامه و الگوریتم اولیه، پایگاه اطلاعاتی نادرست و ناکافی، مستندات و هشدارهای نادرست و ناکافی، به روز نبودن دانش کسب شده، ورودی های نادرستی که افراد ممکن است به عامل هوشمند ارائه دهند و یا به کارگیری برنامه و توانمندی ها در راستای هدفی نادرست (Kingston, 2016, p. 273)؛ اما فراتر از این، قابلیت آسیب رسانی عامل هوشمند می تواند به مراتب بیش از سایر انواع هوش مصنوعی نیز باشد؛ چراکه این عامل ها ضمن یادگیری در سطوح بالاتر، اهدافی را می توانند در نظر بگیرند که به خاطر منافعی، بالقوه امکان ضرر رسانی به دیگران را افزایش می دهد. این اهداف عبارت اند از: محافظت از خود، حفاظت از اهدافی که دارند، تقویت هوش برای بهسازی تصمیم گیری ها و در نهایت کسب حداکثری از منابع به منظور تحقق اهدافشان (Čerka et al., 2015, pp. 382-383)؛ بنابراین نباید پنداشت که با واگذاری فعالیت های انسانی به هوش مصنوعی، مخاطرات ناشی از آن ها به صفر می رسد؛ چنان که همین خطر آفرینی هوش مصنوعی است که در سیاست گذاری های موجود، مقسمی برای طبقه بندی و پیش بینی برخورد حقوقی متناسب با آن قرار گرفته است.

در انتساب رکن مادی رفتارهای آسیب رسان به عامل هوشمند، توجه به تفکیکی مهم و مشترک در مسئولیت کیفری و مدنی، یعنی تفکیک میان مرتکب مادی رفتار و مسئول رفتار ضروری است. به عبارت دیگر، نخست باید دید رفتار از سوی چه کسانی و یا حتی چه چیزهایی امکان ارتکاب دارد و دوم، چه اشخاصی از بابت این رفتار و حسب مورد زیان حاصل از آن مسئول شناخته می شوند. این دو امر گاهی در ادبیات حقوقی و به ویژه در ارتباط با مسئولیت اشخاص حقوقی خلط شده است. تا به امروز مسلم آن است که رفتار، اعم از آن که خطای کیفری یا مدنی باشد، تنها از سوی موجودیتی غیراعتباری قابلیت ارتکاب دارد و اشخاص اعتباری حقوقی، که هیچ گونه موجودیت مادی ندارند، مرتکب مادی رفتار محسوب نمی شوند. همچنین ارتکاب این رفتارها منحصر در رفتارهای انسانی نبوده و حیوانات، اشیا و حتی ذرات در جهان طبیعت - حوادث طبیعی - نیز قابلیت آسیب رسانی به دیگران را دارند؛ لیکن زمانی موجودیت های اخیر پای در عالم حقوق گذاشته و جبران ضرر ناشی از عملکرد آن ها الزامی می نماید که به نحوی به یک رفتار انسانی، اعم از فعل، ترک فعل یا حتی وضعیت و حالت، مستند شوند و در طول آن قرار گیرند. به عبارتی، ضرر ناشی از فعل حیوان، اشیا و حوادث طبیعی مستقلاً موضوع حقوق قرار نمی گیرند؛ بلکه از آنجاکه تحت مالکیت یا وظیفه مراقبت، نگهداری و حفاظت توسط انسان قرار دارند، موجب مسئولیت می شوند (Yazdanian, 2017, p. 384).

1: See: Samuel, P., Jayashree, K., Babu, R., & Vijay, K. (2023). "Artificial Intelligence, Machine Learning, and IoT Architecture to Support Smart Governance." In: AI, IoT, and Blockchain Breakthroughs in E-Governance, IGI global: 95-113.

2: Automated Function

3: Autonomous Action

این امر اما درخصوص هوش مصنوعی نیازمند دقت بیشتری است. نخست آن که هوش مصنوعی از هر قسم و طبقه‌ای که باشد، موجودیتی غیراعتباری است که به‌عنوان مرتکب مادی زیان می‌تواند نقش‌آفرینی کند. دوم تنها آن دسته از هوش مصنوعی که بنابر معیارهای پیش‌گفته، هوش مصنوعی ضعیف خوانده می‌شود و مانند ابزاری در دست انسان است، می‌تواند با همین ضابطه مشابه که درخصوص حیوانات و اشیا بیان شد، موضوع مسئولیت‌های مدنی از قبیل مسئولیت ناشی از فعل شیء، مسئولیت ناشی از کالای معیوب (Pagallo, 2013, p. 116)، مسئولیت محض اشخاص مرتبط با سیستم هوشمند براساس هدف مشترک (Čerka et al., 2015, p. 386) و به‌طورکلی مسئولیت نیابتی قرار گیرد (Čerka et al., 2015, p. 385) و مسئولیت و تحمل ضمانت‌اجرا را بر دوش ارائه‌دهنده، سازنده، کاربر و یا مالک قرار دهد (Rajabi, 2019, pp. 456-458). بنابراین درخصوص عامل هوشمند، که با هویتی مستقل دارای شخصیت قانونی است، به چنین پیش‌شرطی نیاز نیست. باین‌حال، برخی معتقدند دست‌کم در قلمرو کیفری، رابطه سببیت میان ضرر و فعل هوش مصنوعی ضعیف به‌دلیل ارادی‌نبودن رفتار برقرار نمی‌شود و حتی رکن مادی نیز قابلیت انتساب به آن‌ها را ندارد (Claussén Karlsson, 2017, p. 45). این دیدگاه برپایه سنت ارسطویی و در پیوند رفتار مجرمانه با خطای اخلاقی توجیه می‌شود که رابطه سببیت و استناد را برپایه اختیار پی‌ریزی می‌کند (Borhani, 2017, p. 258). این درحالی است که به‌موجب نظریات معاصر در مسئولیت مدنی، هنگامی که ضرر ناشی از فعل زیان‌بار رخ دهد، بین فعل و ضرر رابطه سببیت برقرار می‌شود؛ اعم از آن که رفتار از روی اراده و اختیار انجام شده باشد یا خیر (Badini, 2005, p. 98).

ازاین‌رو، به نظر می‌رسد درخصوص امکان انتساب مادی رفتار و صدمات ناشی از فعل یا ترک فعل عامل دارای هوش مصنوعی به وی، با محدودیت‌هایی که درخصوص اشخاص حقوقی یا فعل حیوان و شیء مطرح می‌شود، روبه‌رو نخواهیم بود؛ چراکه اولاً عامل هوشمند موجودیتی مادی و غیراعتباری است که هرچند ممکن است ساختار غیرروباتیک داشته باشد و فاقد ابزارهای فیزیکی حرکتی باشد و تنها در محیط و فضایی مجازی کنشگری داشته باشد، متکی بر افزارهای اطلاعاتی و ارتباطاتی است و با اثرگذاری ملموس و بی‌واسطه بر محیط، اقداماتش مستقیماً با حواس ما درک‌شدنی است. این درحالی است که در اشخاص حقوقی، چیزی جز انتزاع و اعتبار ما نیست که تصور شخصیتی از آن را برایمان ممکن می‌سازد. همچنین، در این زمینه با محدودیت پیش‌گفته درخصوص اشیا، حیوان و حتی هوش مصنوعی ضعیف در انتساب رفتار به خود عامل هوشمند نیز مواجه نیستیم؛ زیرا منشأ اصلی این امر، که باوجود امکان برقراری رابطه سببیت میان ضرر و فعل حیوان، شی و هوش مصنوعی ضعیف، ضروری است عمل آن‌ها در طول اعمال انسانی قرار گیرد، آن است که درواقع این موجودیت‌ها توانایی و اهلیت پذیرش عواقب ناشی از انتساب مادی این آسیب‌ها را نداشتند و مادامی که فعلشان درنهایت به یک رفتار انسانی متصل نمی‌شد، گویی ضرری است که خارج از عالم حقوق تحقق یافته و فقط در زمره اتفاقات و پیشامدهای طبیعی در عالم خارج است.

اما درخصوص عامل هوشمند که از خودمختاری و عملکرد تا حد زیادی مستقل از انسان برخوردار است و حقوق با اعطای شخصیت قانونی، آن را واجد اهلیت لازم برای تحمل نتایج اعمال خود مستقل از انسان‌ها می‌داند، دیگر التزام به پیش‌شرط ذکرشده برای به‌میدان‌آوردن آسیب وارده از سوی وی به عرصه حقوقی غیرموجه خواهد بود؛ زیرا در جایی علاوه بر انتساب ضرر به فعل یک حیوان یا شیء، نیازمند برقراری رابطه استناد میان آن ضرر و رفتار انسانی بودیم که عاملان بلاواسطه زیان جزء کنشگران به‌رسمیت شناخته‌شده و مخاطبان عالم حقوق نبوده و بدون وجود رابطه‌ای با گروه اخیر، نمی‌توانستند موضوع الزامات حقوقی قرار گیرند؛ اما عامل‌های دارای هوش مصنوعی که خود در کنار انسان به‌منزله موجودیت‌هایی محسوس و مادی و نه اعتباری، دارای شخصیت لازم و حقوق و تکالیف می‌شوند، شخصاً مخاطبان مستقل مقررات حقوقی قرار می‌گیرند و بنابراین، عواقب اعمالشان به‌نحو مستقل و بدون اتکا به سایر کنشگران حقوقی قابل پیگیری است (Čerka et al., 2015, p. 384). در پایان، نباید فراموش کرد که تعمیم رفتارهای ارتكابی عامل دارای هوش مصنوعی به کنشگران دیگر و در نتیجه، توسعه مسئولیت‌های نیابتی برای ارائه‌دهنده یا سازنده عامل هوشمند، اعم از کیفری و مدنی، خود می‌تواند مانعی برای توسعه این دانش و روی‌گردانی متخصصان از ورود به این حوزه شود (Takhshid, 2021, p. 245).

۳. بند سوم: اسناد رکن روانی به عامل دارای هوش مصنوعی

بیش از انتساب رکن مادی رفتار مجرمانه و زیان بار، این رکن روانی در عامل هوشمند است که سخن از مسئولیت این کنشگران را دشوار می‌سازد. به عبارت دیگر، هرچند ایراد ضرر از سوی عامل دارای هوش مصنوعی بدون نیاز به انتساب به فعل یا ترک فعلی انسانی می‌تواند بستر و سبب تحمیل مسئولیت بر وی را فراهم کند، اسناد مؤلفه تقصیر و به‌طور کلی وضعیت‌های ذهنی به انواع هوش مصنوعی محل بحث جدی است.

تقصیر در قلمرو حقوق کیفری، به‌رغم همه تحولاتی که پشت‌سر گذاشته است، همچنان رکنی بنیادین به شمار می‌رود. این مفهوم، از یک‌سو در معنای عام به‌عنوان «رابطه خاص روانی و ذهنی که بین فاعل و جرم برقرار است» (Mirsaeedi, 2007, p. 67)، در قالب رکن روانی جرم پدیدار می‌شود که امکان انتساب جرم را به مرتکب فراهم می‌آورد و از سوی دیگر، در معنای اخلاقی خود، یعنی سرزنش نمود می‌یابد که به تبع آن در اهلیت جزایی، مؤلفه‌های ادراک، تمیز و اختیار را برای مسئولیت کیفری شرط می‌سازد. در مقابل، تقصیر در جریان دگرگونی نظریات در قلمرو مسئولیت مدنی، نه تنها از مفهومی اخلاقی تغییر چهره داد، بلکه حتی در برخی انواع مسئولیت به کلی کنار گذاشته شد. از این رو، به دلیل تمایز بنیادین رکن روانی در مسئولیت مدنی و کیفری ضروری است تبیین هریک را جداگانه در پیش بگیریم.

به‌طور کلی در سیر تحولات مسئولیت مدنی شاهد دو امر هستیم: نخست تحول معیار و مفهوم تقصیر از یک مفهوم اخلاقی مبتنی بر سرزنش به مفهومی اجتماعی با معیاری نوعی؛ و دوم تحولات مبانی مسئولیت مدنی از نظریات مبتنی بر تقصیر تا جایگزینی با مبانی غیرتقصیری و محض که فقط مبتنی بر رابطه سببیت هستند. لزوم وجود تقصیر در برخی از انواع مسئولیت، میراث نظریه قدیمی تقصیر در مسئولیت مدنی بود که سعی داشت این نوع از مسئولیت را با اخلاق پیوند دهد و در پی انقلاب صنعتی، با انتقادهای جدی روبه‌رو شد (Katouzian, 2009, p. 194). با این حال، همچنان مسئولیت مدنی مبتنی بر تقصیر، در قالب نظام‌های حقوقی به‌عنوان اصلی انکارناپذیر تلقی می‌شود و سایر معیارها در راستای تکمیل آن است که در برخی از انواع مسئولیت، پایه و بنیاد قرار می‌گیرد (Qasemzadeh, 2006, p. 342).

در تعریف سنتی و رایج از تقصیر در مسئولیت مدنی، تقصیر عملی است نامشروع که قابل انتساب به مرتکب باشد. این تعریف بر دو عنصر استوار است: نخست عنصر نوعی؛ یعنی نامشروع بودن نفس عمل که در بند پیشین بدان پرداخته شد و دوم عنصر شخصی؛ یعنی قابلیت استناد عمل نامشروع به مرتکب (Katouzian, 2003, pp. 336-338). به موجب این معیار، هرچند ضرورت ندارد مرتکب نتیجه نامشروع را اراده کند، اما باید قوه ادراک و تمیز و به عبارتی قابلیت سرزنش داشته باشد. روشن است که بر مبنای این دیدگاه، مسئولیت شخص حقوقی توجیه‌پذیر نیست؛ چراکه قوه ادراک و تمیز در این اشخاص وجود ندارد و حتی گاهی قدرت مالی لازم نیز از سوی آن برای جبران خسارت مشاهده نمی‌شود (Katouzian, 2003, p. 338). در این رویکرد، عاملی که آگاهی ندارد و قادر نیست رفتار خود را کنترل کند، عملش از روی اراده آزاد نبوده و نمی‌تواند مسئول باشد (Monterossi, 2020, p. 5). در تعاریف دیگر از تقصیر که کم هم نیستند، معیار بر نفس تعهد و عمل زیان‌بار قرار می‌گیرد؛ یعنی تقصیر در معنای تجاوز از تعهدی قانونی یا قراردادی که شخص به عهده داشته است یا تجاوز از تعهدی که وی می‌توانسته آن را بشناسد و اجرا کند.

بنابه برخی تعاریف، که چهره اجتماعی تقصیر در آن‌ها نمود بیشتری می‌یابد، «این رابطه اجتماعی است که حق و تکلیف را به وجود می‌آورد و برای کسی که وظایف ناشی از آن را انجام نداده است، مسئولیت ایجاد می‌کند» (Katouzian, 2003, pp. 339, 342). این تعریف و تعریف دیگری که تقصیر را در تجاوز از رفتاری که لازمه حمایت قانون از دیگران است معنا می‌کند در واقع مبنای مسئولیت مدنی را حمایت از حقوق اشخاص در زندگی اجتماعی در نظر می‌گیرد و همین جاست که گسست تقصیر از اخلاق در مسئولیت مدنی رقم می‌خورد؛ زیرا بر این مبنا، هدف از مسئولیت تنبیه عامل زیان نیست؛ بلکه تأمین هم‌زیستی اجتماعی و حمایت از دیگران است (Katouzian, 2009, p. 208). این دیدگاه‌های اخیر، «معیار تمیز کار خطا از صواب را نوعی و تقصیر را مفهومی اجتماعی بیان می‌کند، نه مفهومی روانی و اخلاقی» (Katouzian, 2003, pp. 343-344). بنابراین، در معیارهای غیر اخلاقی تقصیر، برخلاف مسئولیت کیفری، «نفس تجاوز به معیار عرفی رفتار

مباح در زندگی اجتماعی است» که موجب مسئولیت است (Katouzian, 2003, p. 345)؛ به‌گونه‌ای «مستقل از واردکننده زیان و آنچه در درون وی می‌گذرد» (Badini, 2005, p. 191)؛ ولو آن‌که همراه با هیچ‌گونه سهل‌انگاری نباشد.

بنابراین از آنجا که هدف در مسئولیت مدنی، هماهنگی میان آزادی و تکالیف و امتیازهای زندگی اجتماعی است، تقصیر مفهومی نوعی و اجتماعی دارد و صرفاً باید دید رفتاری که از مرتکب سرزده است، قطع نظر از شخصیت و اهلیت وی، از نظر نوعی و عرفی با رفتار انسان متعارف تعارض دارد یا خیر، چه صغیر و مجنون باشد، چه یا بالغ و متخصص. این دیدگاه حذافصل میان دو نظریه تقصیر و خطر در مبنای مسئولیت مدنی است (Katouzian, 2003, pp. 350-351). همین معیار نوعی و مادی در تقصیر بود که راه را برای مسئولیت مدنی اشخاص حقوقی گشود؛ بدون آن‌که نیاز به عنصر معنوی تقصیر و سرزنش اخلاقی در آن‌ها باشد (Qasemzadeh, 2006, p. 127). بنابراین، با تغییر معیار تقصیر در مسئولیت‌های مبتنی بر تقصیر از «تقصیر در عامل» به «تقصیر در عمل» و طرح مسئولیت مدنی ناشی از تقصیر به‌عنوان «مسئولیت بدون سرزنش» (Qasemzadeh, 2006, p. 214)، دشواری‌هایی که در مسئولیت مدنی برای عامل دارای هوش مصنوعی می‌تواند مطرح شود، کاهش چشمگیری می‌یابد؛ اگرچه عده‌ای تغییر ملاک تقصیر از عامل به عمل را الزاماً موجب جدایی مسئولیت مدنی از سرزنش‌پذیری اخلاقی نمی‌دانند (Ghiasvand, 2021, p. 240).

همچنین نظریات مبتنی بر نگرش اقتصادی نیز مبنای مسئولیت مدنی عامل هوشمند را تقویت می‌کند. درجایی که هدف مسئولیت مدنی بنابر نگرش اخیر، کارایی اقتصادی قرار می‌گیرد و نه صرف جبران خسارت، عامل زیان موجودی دارای عقلانیت در نظر گرفته می‌شود که بر هزینه فعالیت‌ها تمرکز می‌کند و واداشته می‌شود تا مجموع هزینه‌های ناشی از رفتارهای زیان‌بار و حوادث و هزینه‌های پیشگیری از آن رفتارها و حوادث را به حداقل برساند. این امر از دو طریق محقق می‌شود: نخست انجام‌دادن فعالیت‌های مستعد زیان و دوم افزایش مراقبت‌های هنگام انجام این فعالیت‌ها (Qasemzadeh, 2006, pp. 427-428). درحقیقت، این نظریات بر میزان هزینه رفتارهای پیشگیرانه و احتیاطی و میزان هزینه‌ای که از انجام رفتارهای زیان‌بار به بار می‌آید، تأکید می‌کنند و عامل زیان را به‌عنوان موجودی دارای عقلانیت به این سمت سوق می‌دهند که اگر هزینه نهایی رفتارهای زیان‌بار، بیش از هزینه‌های پیشگیری و احتیاطی (در مسئولیت مبتنی بر تقصیر) یا منفعت ناشی از انجام فعالیت‌های بالقوه زیان‌بار باشد (مسئولیت محض)، با قبول مسئولیت، متعهد به پرداخت آن هزینه‌های نهایی شود (Qasemzadeh, 2006, p. 444). بر مبنای بازدارندگی اقتصادی در توجیه مسئولیت مدنی، به نظر می‌رسد اتفاقاً چنین حسابگری و محاسبات عقلانی‌ای را می‌توان با توجه به توانمندی‌ها و الگوریتم‌های دقیق محاسباتی هوش مصنوعی و داده‌های گسترده‌ای که از محیط پیرامون و به طرق مختلف کسب می‌کنند، با دقت بیشتری در عامل‌های هوشمند ملاحظه کرد و انتظار بیشتری داشت. بهره‌گیری از الگوریتم‌ها در تقویت مسئولیت‌پذیری تاجایی است که برخی معتقدند به‌کارگیری معیار معقول و متعارف بودن الگوریتم‌ها در هوش مصنوعی، حتی ما را از پیچیدگی‌های احراز عملکرد ذهنی و فرایند تصمیم‌گیری عامل زیان در انتساب مسئولیت بی‌نیاز می‌سازد (Zakerinia & Gholampour, 2024, p. 162).

در نقطه مقابل، مسئولیت کیفری اصولاً بر مؤلفه‌هایی شخصی قرار می‌گیرد و عنصر تقصیر را در مفهوم اخلاقی آن یعنی سرزنش‌پذیری ملاک می‌داند. براین‌اساس سرزنش عامل یک رفتار زمانی معنا می‌یابد که از نظر اخلاقی دارای ادراک و تمیز شناخته شود؛ امری که محل اختلاف پژوهشگران و اندیشمندان حوزه هوش مصنوعی و دانش‌های مرتبط با آن است^۱. به‌بیان دیگر، در جایی که تقصیر، خواه عمدی و خواه غیرعمدی، خواه سنگین و فاحش و خواه سبک، مؤلفه و رکن رکن مسئولیت حقوقی به شمار می‌آید، مادامی که نتوان اهلیت و قابلیت برخورداری شخص از مسئولیت اخلاقی را احراز کرد، نمی‌توان وی را از نظر حقوقی نیز مسئول دانست، حتی اگر تقصیر معیاری نوعی و اجتماعی یابد (Qasemzadeh, 2006, p. 127).

این تقصیر اخلاقی در مسئولیت کیفری، به‌نحو پیشینی، وجود آگاهی و اراده را در مرتکب ایجاب می‌کند. مقصود از آگاهی، داشتن وضعیت ذهنی سالمی است که امکان شناخت شخص نسبت به خود و محیط پیرامون و انطباق یا عدم انطباق رفتار خود با واقعیت‌های قانونی

1: See: Sullins, J. P. (2006). When is a robot a moral agent. International Review of Information Ethics, 6(2006), 23-30.

را فراهم می‌آورد (Mosavi Mojab, 2009, p. 115). به بیان دیگر، شخصی از منظر اخلاقی سرزنش‌پذیر است که بالقوه توانایی آگاهی‌یافتن از موضوع، شرایط، محیط و ضوابط ناظر بر رفتار خود را داشته باشد. به عقیده متخصصان، مراحل آگاهی در عامل هوشمند با دریافت اطلاعات از محیط و ایجاد تصویر از داده‌ها و پردازش آن‌ها نزد وی تکمیل می‌شود (Hallevy, 2019, p. 95)؛ تاجایی که می‌تواند در محاسبه احتمالات و کنشگری متناسب با آن‌ها، بهتر از عامل انسانی عمل کند (Stone et al., 2016, pp. 12-13). هرچند از نظر کسانی که قابلیت سرزنش اخلاقی را علاوه بر آگاهی و اراده، مستلزم وجود خودآگاهی، یعنی آگاهی از موجودیت، باورها و اراده خویشتن می‌دانند، بهره‌مندی عامل هوشمند از عنصر روانی و اسناد مؤلفه‌های آن به او با مانع روبه‌رو می‌شود (Gless et al., 2016, p. 416)؛ به‌ویژه آن‌که نیل به اهداف مجازات را نیز درخصوص این عامل‌ها ناممکن می‌دانند (Amirian Farsani & Hosseini, 2024, p. 42).

در باب بهره‌مندی از اراده نیز علاوه بر دلایلی که در بخش‌های پیشین درخصوص عملکرد فنی مستقل عامل هوشمند عنوان شد، برخی حتی تا جایی پیش رفته‌اند که عامل هوشمند را یگانه عامل مختار در نظر می‌گیرند؛ چراکه برخلاف انسان، این عامل هوشمند است که به‌واقع می‌تواند نشئت‌گرفته از دلایل و اسباب اندیشیده شده و صرفاً بنابر فرضیات منطقی محض عمل کند و لذا تنها اوست که آزاد است (Nadeau, 2006, p. 241-248).^۱ بالاین حال، تضارب آرا از نفی فلسفی امکان آگاهی، اراده و قصدمندی در عامل دارای هوش مصنوعی تا برخورداری ناچیز آن‌ها از این مؤلفه‌ها، رو به کاهش نگذاشته و همین امر، راه پیاده‌سازی مسئولیت کیفری عامل هوشمند را هرچند ناممکن نمی‌گرداند، اما به‌مراتب دشوار می‌سازد. در این زمینه، حتی تحولات ناظر بر مسئولیت کیفری اشخاص حقوقی نیز نمی‌تواند ما را از لزوم وجود اهلیت جزایی برای عامل هوشمند بی‌نیاز کند؛ زیرا در مسئولیت اشخاص حقوقی نیز یا اهلیت شخص حقوقی از کارکنان، مدیران و مغزهای متفکر آن به عاریت گرفته می‌شود یا نوعی اراده جمعی و فرهنگ سازمانی جایگزین اراده فردی نیابتی شده و از این طریق مسئولیت کیفری این اشخاص توجیه می‌شود (Goldozian & Golriz, 2018, pp. 216-217). به همین دلیل، امروزه نظام‌های حقوقی داخلی و بین‌المللی با احتیاط کامل گام برداشته و همچنان با نگاه ابزارگونه به هوش مصنوعی و تلقی آن به‌عنوان یک دارایی، از شناسایی مسئولیت شخصی برای این عامل‌ها، به‌ویژه در قلمرو کیفری، سر‌باز می‌زنند (Kan, 2024, p. 289). اگرچه باوجود چالش‌های اثبات شیوه و چگونگی عملکرد آگاهی، اراده و قصد و تقصیر در عامل هوشمند، حتی از منظر فنی، که تحت‌عنوان «شفافیت و توضیح‌پذیری هوش مصنوعی»^۲ در مباحث علمی دنبال می‌شود، شاید تقویت این قابلیت‌ها در کنار عبور از دیدگاه‌های سنتی و استقبال از نظریه‌های نو، بتواند پذیرش مسئولیت کیفری برای این عامل‌ها را با مقبولیت بیشتری همراه کند.

درمجموع و با توجه به مطالب فوق، گفتنی است مسئولیت مدنی عامل هوشمند به‌مراتب با پیچیدگی کمتری به نسبت شناسایی مسئولیت کیفری برای آن‌ها مواجه است؛ چراکه نه تنها عنصر تقصیر، که یک پای آن در وضعیت ذهنی و قوه ادراکی شخص مقصر قرار دارد، در مسئولیت مدنی همواره شرط ضروری نیست، بلکه افزون بر آن، معیار تقصیر در همین موارد نیز بر مبنای بی‌احتیاطی و بی‌مبالاتی و به‌طور کلی سرزنش‌پذیری اخلاقی قرار نمی‌گیرد تا ما را با تنگناهای مباحث مربوط به تفکر و وضعیت‌های ذهنی در هوش مصنوعی روبه‌رو سازد. این درحالی است که در قلمروی کیفری پس از شناسایی شخصیت قانونی برای عامل دارای هوش مصنوعی، احراز ظرفیت‌های اخلاقی این کنشگران نیز ضروری است؛ زیرا برخورداری از مسئولیت اخلاقی و قوای ادراکی و شناختی، شرط لازمه مسئولیت از نظرگاه حقوق کیفری است که امکان‌سنجی تفصیلی آن از حوصله این نوشتار خارج است.^۳

1: Sullins, J. P. (2006). When is a robot a moral agent. *International Review of Information Ethics*, 6(2006), 27.

2: Transparency and Explainability of AI. See: Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., and Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, 159, 107197.

۳. ر.ک: خادمی شورمستی، یاسمن (۱۳۹۸)، امکان‌سنجی مسئولیت کیفری و مجازات هوش مصنوعی، پایان‌نامه کارشناسی ارشد، دانشکده حقوق و علوم سیاسی دانشگاه تهران.

نتیجه‌گیری

امروزه با عنایت به گسترش تحقیقات و پیشرفت‌های سریع در ساخت سیستم‌هایی که ادعا می‌شود قدم به قدم به شباهت در شیوه و ابزارهای تفکر و رفتار انسانی نزدیک می‌شوند، بیش از هر زمان دیگری لزوم تعامل نزدیک هوش مصنوعی و حقوق احساس می‌شود. از میان ارتباطات متقابلی که حقوق و هوش مصنوعی دارند، قرارگیری هوش مصنوعی به منزله مخاطب جدید حقوق و تکالیف در یک نظام حقوقی بیش از همه جلب توجه می‌کند. عامل‌های دارای هوش مصنوعی به عنوان سطحی از سیستم‌های هوشمند، که پیش‌بینی می‌شود بتوانند با یادگیری و کسب تجربه از محیط، به نحوی خودمختار برای کسب منافع و اهداف شخصیشان عمل کنند، این قابلیت را دارند که در راستای تحقق این اهدافشان آسیب‌هایی را به موجودات پیرامون خود وارد سازند. همین امر، لزوم شناسایی شخص یا اشخاصی که بتوانند مسئولیت این آسیب و زیان‌ها را، خواه در قلمروی کیفری و خواه در قلمرو مدنی، عهده‌دار شوند ضروری می‌گردد.

در این پژوهش، با نگاهی به مبانی مسئولیت کیفری و مدنی، این نتیجه حاصل شد که پیش‌شرط ضروری تحمیل مسئولیت حقوقی بر عامل‌های هوشمند به نحو مستقل از عامل‌های انسانی، اعطای شخصیت قانونی به آن‌ها به معنای برخورداری از مجموعه‌ای امتیازات و استحقاقات قانونی به منزله مبنای حق و تکلیف است. در مسئولیت مدنی، با توجه به تغییر معیارهای تقصیر از ملاک اخلاقی به ملاک اجتماعی از یک سو و حذف مؤلفه تقصیر در برخی از انواع مسئولیت، چالش‌هایی که پیرامون امکان سرزنش‌پذیری اخلاقی عامل‌های دارای هوش مصنوعی مطرح می‌شود، با پاسخ‌های روشن‌تر و دشواری کمتری برای این عامل‌ها روبه‌روست. در مقابل، مسئولیت کیفری به دلیل پیوند همیشگی خود با مسئولیت اخلاقی، ممکن است قدری در اقناع‌سازی متخصصان برای شناسایی این نوع مسئولیت برای عامل هوشمند با مشکل مواجه شود، اما رهایی از نگاه سنتی به آن و پذیرش نظرات نو مسیر را هموار می‌کند.

در مجموع، به نظر می‌آید آنچه موجب تردید در خصوص مسئولیت شخصی و مستقیم عامل هوشمند می‌شود، نه نفس امکان انتساب رکن مادی و نه حتی تقصیرپذیری این عامل‌ها، بلکه نوع نگرش حقوق و نظام حقوقی به این کنشگران نوظهور به عنوان شخصیت‌های قانونی است؛ بنابراین اگر نظام حقوقی شخصیت قانونی همراه با حقوق و تکالیف و تعهدات معین و متناسب را برای عامل هوشمند به رسمیت بشناسد، آن زمان است که این عامل به موجب کنش‌ها و رفتارهایی که انجام می‌دهد و نتایج و عواقب آن‌ها، مسئول است؛ اگرچه شناسایی شخصیت قانونی از موضوعاتی است که بیش از مباحث حقوقی با موضوعات فلسفی، اجتماعی و اخلاقی پیرامون عامل هوشمند است که پیوند می‌یابد.

بدین ترتیب و با وجود تمرکز ادبیات حقوقی کنونی بر مسئولیت‌پذیری عامل انسانی، به جای عامل هوشمند، نباید فراموش کرد که اگر حقوق با تکیه بر وضع موجود، صرفاً نظاره‌گر تحولات جامعه قرار گیرد، آن زمان به دلیل آماده‌نبودن در مواجهه با پدیده‌هایی که به سرعت و یک‌باره ممکن است در جامعه متولد شوند، نه تنها کارایی خود را در حکمرانی و حفظ نظم از دست خواهد داد، بلکه جبران این عقب‌ماندگی حقوقی، نظام حقوقی و جامعه را متحمل صرف هزینه‌های بعضاً جبران‌ناپذیری خواهد کرد.

منابع

- ابوذری، مهرانوش (۱۴۰۱). حقوق و هوش مصنوعی. چ ۲. تهران: میزان.
- امیریان فارسانی، امین و حسینی، سیدمحمد (۱۴۰۲)، چالش‌ها و موانع مسئولیت کیفری در ربات‌های با قابلیت هوش مصنوعی، تمدن حقوقی، ۶(۱۸)، ۳۹-۵۴. <https://doi.org/10.22034/lc.2024.449638.1458>
- بادینی، حسن (۱۳۸۴). فلسفه مسئولیت مدنی. چ ۱. تهران: شرکت سهامی انتشار.
- برهانی، محسن (۱۳۹۶). اخلاق و حقوق کیفری. چ ۳. تهران: پژوهشگاه فرهنگ و اندیشه اسلامی.
- تخشید، زهرا (۱۴۰۰). مقدمه‌ای بر چالش‌های هوش مصنوعی در حوزه مسئولیت مدنی. حقوق خصوصی، ۱۸(۱)، ۲۲۷-۲۵۰. <https://doi.org/10.22059/jolt.2021.319529.1006965>
- تگمارک، مکس (۱۳۹۷). زندگی ۰/۳: انسان بودن در عصر هوش مصنوعی. ترجمه میثم محمدامینی. چ ۱. تهران: نشر نو.
- حکمت‌نیا، محمود، محمدی، مرتضی و واثقی، محسن (۱۳۹۸). مسئولیت مدنی ناشی از تولید ربات‌های مبتنی بر هوش مصنوعی خودمختار. حقوق اسلامی، ۱۵(۶۰)، ۲۵۸-۲۳۱. [لینک](#)
- خادمی شورمستی، یاسمن (۱۳۹۸). امکان‌سنجی مسئولیت کیفری و مجازات هوش مصنوعی. پایان‌نامه کارشناسی ارشد. دانشکده حقوق و علوم سیاسی دانشگاه تهران، تهران.
- ذاکری‌نیا، حانیه و غلامپور، زهرا (۱۴۰۳). الگوریتم‌های معقول و متعارف و تقویت نظریه قابلیت انتساب مسئولیت مدنی هوش مصنوعی. حقوق فناوری‌های نوین، ۵(۹)، ۱۵۵-۱۶۸. <https://doi.org/10.22133/mtlj.2023.406159.1224>
- راسل، استوارت جی و نوروینگ، پیتر (۱۳۹۰). هوش مصنوعی رهیافتی نوین. ترجمه رامین رهنمون. چ ۴. تهران: ناقوس.
- رجبی، عبدالله (۱۳۹۸). ضمان در هوش مصنوعی. مطالعات حقوق تطبیقی، ۱۰(۲)، ۴۴۹-۴۶۶. <https://doi.org/10.22059/jcl.2019.274782.633787>
- رهنمون، رامین (۱۳۸۴). داستان شیرین هوش مصنوعی. چ ۱. تهران: ارسا.
- طاقانی، عبدالعظیم (۱۳۹۶). هوش مصنوعی. چ ۱. مشهد: قلم امامت.
- غیاثوند، مهدی (۱۴۰۰). بررسی وابستگی «مسئولیت مدنی مبتنی بر تقصیر در پارادایم حق و عدالت» به متغیرهای فلسفی. حقوق اسلامی، ۱۸(۷۰)، ۲۳۱-۲۵۱. [لینک](#)
- قاسم‌زاده، سیدمرتضی (۱۳۸۵). مبانی مسئولیت مدنی. چ ۴. تهران: میزان.
- کاتوزیان، ناصر (۱۳۸۲). الزام‌های خارج از قرارداد: ضمان قهری. چ ۳. تهران: دانشگاه تهران.
- کاتوزیان، ناصر (۱۳۸۸). تحول مفهوم تقصیر در حقوق مسئولیت مدنی. مطالعات حقوق خصوصی، ۳۹(۱)، ۱۸۹-۲۱۴. [لینک](#)
- گلدوزیان، ایرج و گلریز، امین (۱۳۹۷). اهلیت جنایی اشخاص حقوقی. مجله حقوقی دادگستری، ۸۲(۱۰۲)، ۲۰۹-۲۳۳. <https://doi.org/10.22106/jlj.2018.32739>
- موسوی مجاب، سیددرید (۱۳۸۸). نقش اراده در مسئولیت کیفری بارویکردی به نظام حقوقی ایران. چ ۱. تهران: بهنامی.
- میرسعیدی، منصور (۱۳۸۶). مسئولیت کیفری. چ ۱. چ ۲. تهران: میزان.
- هالوی، گابریل (۱۳۹۸). مسئولیت کیفری ربات‌ها. ترجمه فرهاد شاهیده و طاهره قنولو. چ ۱. تهران: میزان.
- یزدانیان، علیرضا (۱۳۹۶). طرح قاعده مسئولیت مدنی ناشی از اشیا در حقوق ایران و فرانسه. مطالعات حقوق خصوصی، ۴۷(۲)، ۳۷۹-۳۷۹. <https://doi.org/10.22059/jlq.2017.62621.3997>

- Asaro, P. M. (2007). Robots and responsibility from a legal perspective. *Proceedings of the IEEE*, 4(14), 20-24. [Link](#)
- Atabekov, A., & Yastrebov, O. (2018). Legal status of artificial intelligence across countries: Legislation on the move. *European Research Studies Journal*, 21(4), 773-782. <https://doi.org/10.35808/ersj/1245>
- Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkänen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, 159, 107197. <https://doi.org/10.1016/j.infsof.2023.107197>
- Bertolini, A., Salvini, P., Pagliai, T., Morachioli, A., Acerbi, G., Trieste, L., Cavallo, F., Turchetti, G. and Dario, P. (2016). On robots and insurance. *International Journal of Social Robotics*, 8, 381-391. <https://doi.org/10.1007/s12369-016-0345-z>
- Čerka, P., Grigienė, J., & Širbikytė, G. (2015). Liability for damages caused by artificial intelligence. *Computer Law & Security Review*, 31(3), 376-389. <https://doi.org/10.1016/j.clsr.2015.03.008>
- Chakrabarty, I., & Baral, A. (2023). Artificial Intelligence and Personhood in the 21st Century. *Indian J. Integrated Rsch. L.*, 3, 1. [Link](#)
- Chopra, S., & White, L. F. (2011). *A legal theory for autonomous artificial agents*. University of Michigan Press. <https://doi.org/10.3998/mpub.356801>
- Claussén Karlsson, M. (2017). Artificial intelligence and the external element of the crime: An analysis of the liability problem. Dissertation. Örebro University. [Link](#)
- Dyschkant, A. (2015). Legal personhood: how we are getting it wrong. *U. Ill. L. Rev.*, 2075. [Link](#)
- European Commission. (2022, September 28). Proposal for a directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive). [Link](#)
- European Parliament. (2017, February 16). Civil law rules on robotics. [Link](#)
- European Parliament. (2020, October 20). Civil liability regime for artificial intelligence. [Link](#)
- European Parliament. (2024, June 13). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending regulations (EC) no 300/2008, (EU) no 167/2013, (EU) no 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). [Link](#)
- Gless, S., Silverman, E., & Weigend, T. (2016). If robots cause harm, who is to blame? Self-driving cars and criminal liability. *New Criminal Law Review*, 19(3), 412-436. <https://doi.org/10.1525/nclr.2016.19.3.412>
- Kan, C. H. (2024). Criminal liability of artificial intelligence from the perspective of criminal law: An evaluation in the context of the general theory of crime and fundamental principles. *International Journal of Eurasia Social Sciences (IJOESS)*, 15(55), 276-313. <https://doi.org/10.35826/ijoess.4434>

- Kingston, J. K. (2016, November). Artificial intelligence and legal liability. In *International conference on innovative techniques and applications of artificial intelligence* (pp. 269-279). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-47175-4_20
- Kudina, O., & van de Poel, I. (2024). A sociotechnical system perspective on AI. *Minds and Machines*. Springer, 34(3), 21. <https://doi.org/10.1007/s11023-024-09680-2>
- Kurki, V. A. J., & Pietrzykowski, T. (2017). Legal personhood: Animals, artificial intelligence and the unborn. Springer. [Link](#)
- Michalski, R. (2018). How to sue a robot. *Utah Law Review*. 2018(5), 1021-1071. [Link](#)
- Monterossi, M. W. (2020). Liability for the fact of autonomous artificial intelligence agents. Things, agencies and legal actors. *Global Jurist*, 20(3), 20190054. <https://doi.org/10.1515/gj-2019-0054>
- Nadeau, J. E. (2006). Only androids can be ethical. *Thinking about android epistemology*, 241-248. [Link](#)
- Novelli, C. (2023). Legal personhood for the integration of AI systems in the social context: A study hypothesis. *AI & SOCIETY*, 38(4), 1347-1359. <https://doi.org/10.1007/s00146-021-01384-w>
- Pagallo, U. (2013). The laws of robots: Crimes, contracts, and torts. Springer. <https://doi.org/10.1007/978-94-007-6564-1>
- Samuel, P., Jayashree, K., Babu, R., & Vijay, K. (2023). Artificial intelligence, machine learning, and IoT architecture to support smart governance. In *AI, iot, and blockchain breakthroughs in E-Governance* (pp. 95-113). IGI global. <https://doi.org/10.4018/978-1-6684-7697-0.ch007>
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424. <https://doi.org/10.1017/S0140525X00005756>
- Solum, L. B. (1992). Legal personhood for artificial intelligences. *North Carolina Law Review*, 70(4), 1231-1288. [Link](#)
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A.L., Shah, J., Tambe, M., & Teller, A. (2016). Artificial intelligence and life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel. <https://doi.org/10.48550/arXiv.2211.06318>
- Sullins, J. P. (2006). When is a robot a moral agent?. *International Review of Information Ethics*, 6(2006), 23-30. [Link](#)
- Tasioulas, J. (2019). First steps towards an ethics of robots and artificial intelligence. *Journal of Practical Ethics*, 7(1), 49-83. [Link](#)

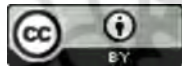
[In Persian]

- Abouzari, M. (2022). *Law and artificial intelligence*. 1st Pub, Tehran: Mizan.
- Amirian Farsani, A. and Hosseini, S. M. (2024). Challenges and Obstacles of Criminal Liability in Robots with Artificial Intelligence Capabilities. *Fares Law Research*, 6(18), 39-54.
<https://doi.org/10.22034/lc.2024.449638.1458>
- Badini, H. (2005). *Philosophy of civil liability*. 1st Pub, Tehran: Enteshar Co.
- Borhani, M. (2017). *Ethics and criminal law*. 3rd Pub, Tehran: Research Institute for Islamic Culture and Thought.
- Ghiasvand, M. (2021). Fault-based civil liability and some fundamental philosophical debates. *Islamic Law*, 18(70), 231-251. [Link](#)
- Goldozian, I. and Golriz, A. (2018). Criminal capacity of legal persons. *The Judiciary Law Journal*, 82(102), 209-233. <https://doi.org/10.22106/jlj.2018.32739>
- Hallevy, G. (2019). *When robots kill*. Translated by Farhad Shahideh & Tahereh Ghavanloo. 1st Pub, Tehran: Mizan.
- Hekmatnia, M., Mohammadi, M. and Vaseghi, M. (2019). Civil liability for damages caused by robots based on autonomous artificial intelligence. *Islamic Law*, 16(60), 231-258. [Link](#)
- Katouzian, N. (2003). *Iranian civil law: A comparative study extra contractual obligations*. 3rd Pub, Tehran: University of Tehran.
- Katouzian, N. (2009). The evolutionary conception of fault in tort law. *Law Quarterly*, 39(1), 189-214. [Link](#)
- Khademi Shoormasti, Y. (2019). Feasibility of criminal liability and punishment of artificial intelligence. *Master's Thesis*. Tehran: University of Tehran.
- Mirsaeedi, S.M. (2007). *Criminal responsibility*. 1st Pub, Vol. 1, Tehran: Mizan.
- Mosavi Mojab, S.D. (2009). *The role of will in criminal liability*. 1st Pub, Tehran: Behnami.
- Qasemzadeh, S.M. (2006). *The foundations of civil liability*. 4th Pub, Tehran: Mizan.
- Rahnamoon, R. (2005). *The sweet story of artificial intelligence*. 1st Pub, Tehran: Arsaa.
- Rajabi, A. (2019). Liability of artificial intelligence; the reflection of developments in the liability rules. *Comparative Law Review*, 10(2), 449-466. <https://doi.org/10.22059/jcl.2019.274782.633787>
- Russell, S. J., and Norvig, P. (2011). *Artificial intelligence: A modern approach*. Translated by Ramin Rahnamoon. 4th Pub, Tehran: Naghoos.
- Taghani, A. (2017). *Artificial intelligence*. 1st Pub, Mashhad: Qalam Imamat.
- Takhshid, Z. (2021). An introductory study on the challenges of artificial intelligence in tort law. *Private Law*, 18(1), 227-250. <https://doi.org/10.22059/jolt.2021.319529.1006965>

- Tegmark, M. (2018). Life 3.0: Being human in the age of artificial intelligence. *Translated by Meysam MohammadAmini*. 1st Pub, Tehran: Nashrenow.
- Yazdanian, A. (2017). The design of the rule of civil liability arising from objects in iran and french law. *Law quarterly*, 47(2), 379-397. <https://doi.org/10.22059/jlq.2017.62621>
- Zakerinia, H. and Gholampour, Z. (2024). Reasonable algorithms and strengthening the "opposability" theory on the civil liability of artificial intelligence. *Modern Technologies Law*, 5(9), 155-168. <https://doi.org/10.22133/mtlj.2023.406159.1224>

COPYRIGHTS

©2026 by the authors. Published by University of Science and Culture. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>



پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی