



Journal of Philosophical Investigations



University of Tabriz

Inclusivity in Artificial Intelligence from a Philosophical-Legal Perspective

Seyed Mohammad Hassan Momeni¹ | Maryam Parvizi²

1. Corresponding Author, Assistant Professor, Department of Jurisprudence and Law, Allameh Tabataba'i University, Tehran, Iran. E-mail: dr.smhmomeni@atu.ac.ir

2. PhD in Comparative Philosophy. Allameh Tabataba'i University, Tehran, Iran. E-mail: maryam_parvizi@atu.ac.ir

Article Info

Article type:

Research Article

Article history:

Received 08 November 2025

Received in revised form 10
January 2026

Accepted 14 January 2026

Published online 18 July 2026

Keywords:

Philosophy of Artificial
Intelligence, Jurisprudence and
Law of Artificial Intelligence,
Citizen, Individuals, Persons,
Learning.

ABSTRACT

The rapid development of AI has drawn intense attention to its philosophical and legal implications, especially regarding discrimination and bias. A fundamental weakness in this discourse is the emphasis on the concept of “citizen,” which is an exclusive category that does not encompass all stakeholders. Achieving true inclusiveness in AI systems requires a fundamental shift in terminology from “citizen” to a more inclusive concept. The title “individual” is a common alternative, but the legal term “person” is more inclusive. Because it encompasses both natural and legal persons. This research emphasizes, using the method of conceptual and critical analysis within the framework of philosophical-ethical, jurisprudential-legal thought and with library tools, that protecting this inclusiveness is not only a moral necessity, but also a legal obligation, and examines the foundations of Imami jurisprudence as a Sharia law perspective to support and explain the person-centered approach. By combining philosophical-ethical, and jurisprudential-legal arguments, this article has provided a more robust framework for the design and governance of pervasive artificial intelligence systems.

Cite this article: Momeni, M. H. & Parvizi, M. (2026). Inclusivity in Artificial Intelligence from a Philosophical-Legal Perspective. *Journal of Philosophical Investigations*, 20 (55), 221-242. <https://doi.org/10.22034/jpiut.2026.70131.4323>



© The Author(s).

Publisher: University of Tabriz.

Extended Abstract

The rapid growth of artificial intelligence has been accompanied by special attention to its philosophical-ethical and legal issues. The increasing publication of various ethical guidelines and efforts to legislate in this area demonstrate the importance of such issues for researchers.

Trifacial intelligence has on society, one of the most important issues is the issue of stakeholders. In the research literature, both in the field of ethics and moral philosophy and in the field of law, human stakeholders of artificial intelligence systems have sometimes been examined as citizens and sometimes as individuals. From a legal perspective, inclusiveness brings different functions and rights to stakeholders. Undoubtedly, in our legal system, the issue of jurisprudential approach and foundations is one of the central functions. Therefore, inclusiveness should be explored from a jurisprudential-legal perspective in addition to a philosophical-ethical perspective. This article attempts to introduce inclusiveness in the strict sense of the word, which necessarily includes philosophical-ethical and jurisprudential-legal aspects, in artificial intelligence systems. It is worth noting that the jurisprudential aspects discussed concern Imamiyya jurisprudence as one of the main forms of governance in Iran and an important foundation of the domestic legal system. Therefore, wherever the jurisprudential-legal aspect is mentioned in this article, it refers to Imamiyya jurisprudence.

The purpose of the present study is to examine philosophical-ethical and jurisprudential-legal inclusiveness in the field of artificial intelligence, and in this regard, attention has also been paid to the socio-technical systems approach. Although inclusiveness has been recommended many times in the research literature, the hypothesis of this study is that, whether from a philosophical-ethical perspective or a jurisprudential-legal perspective, and especially based on the socio-technical systems approach, the ethical and legal performance of artificial intelligence systems requires not focusing on citizen stakeholders and paying attention to all stakeholders, regardless of their citizenship status. In order to achieve the main goal, the present article has been compiled in three parts. Part one, the socio-technical systems approach and the necessity of inclusiveness. Part two, inclusiveness in socio-technical systems as a "basic philosophy" approach in relation to the ethics of artificial intelligence. Part three, inclusiveness in the field of artificial intelligence from a jurisprudential-legal perspective, taking into account the socio-technical systems approach. The necessity of the present study is because no independent research has yet addressed the need for inclusiveness in artificial intelligence, both from a philosophical-ethical and a jurisprudential-legal perspective, especially with regard to the socio-technical system. Artificial intelligence systems, as a socio-technical system, have additional fundamental components due to the different nature of artificial intelligence. The three characteristics of autonomy, interactivity, and adaptability are the main differences between artificial intelligence and traditional technologies. It should be noted that artificial intelligence, like traditional socio-technical systems, has the ability to learn from the environment and adapt to it. However, the difference is that in traditional socio-technical systems this is achieved with human agents, but artificial intelligence systems have technical elements that enable direct learning from the environment. Two additional constituent components in artificial intelligence systems are artificial agents and technical rules. Artificial agents are non-human and passive agents that can act autonomously. And adapt to their environment inputs. Technical rules create and regulate interactions between artificial agents

and other technologies. The role of these rules in artificial intelligence systems can be considered similar to the role of social institutions in traditional socio-technical systems. Inclusivity can act as a counter-bias. AI should not be limited to specific classified information and bound to a specific orientation. For example, race, color, and nationality should not be influential in programs that are implemented and managed by AI tools. The outcome of the present research in two extremely necessary dimensions of ethics and law at the international community level in order to pay attention to inclusiveness in artificial intelligence is based on the meaning that it is necessary to achieve awareness in order to observe the general aspect and inclusiveness, especially based on socio-technical systems, which provide grounds for bias and discrimination, violation of privacy, and ignoring experiential and efficient ethics in a world immersed in artificial intelligence and full of intelligent robots. To achieve a comprehensive and inclusive ethical system, it is necessary to form a generalizable and acceptable ethical system based on various philosophies. In the legal aspect, attention to the principles and rules of Imami jurisprudence in the field of domestic law can combine the feature of inclusiveness with stronger legal and customary foundations, and thus, from a philosophical-moral and jurisprudential-legal perspective, the feature of inclusiveness can be sought in a comprehensive program in artificial intelligence systems. To achieve this extremely important goal, one of the necessary steps is to go beyond the concept of citizenship and pay attention to replacing it with a universal factor. Because non-citizens who are interested are also under the control of smart machines and the world of artificial intelligence, and relying solely on the element of citizenship in such a complex and sensitive world is not only not a guarantee for human rights and fundamental human rights, but it is also a basis for all kinds of discrimination and bias, violation of individual privacy, and dozens of other immoral and anti-legal principles, let alone the identity of socio-technical systems that, regardless of these consequences for the very inclusivity of the social-cultural value, due to the nature of technology, and inclusivity should be considered from this perspective. Therefore, perhaps the title of "individual" instead of citizen could be a suggested option in this regard. But it does not seem to be a sufficient and efficient option. Because its inclusion seems a bit difficult compared to legal entities. Therefore, the option of "person and persons" can be a suitable suggestion due to its inclusion of natural and legal persons at the general level, including citizens and others.



شمول‌گرایی در هوش مصنوعی از نگاه فلسفی - حقوقی

محمدحسن مومنی^۱ | مریم پرویزی^۲

۱. نویسنده مسئول، استادیار گروه فقه و حقوق، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: dr.smhmomeni@atu.ac.ir

۲. دانش‌آموخته دکتری فلسفه تطبیقی، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: maryam_parvizi@atu.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی تاریخ دریافت: ۱۴۰۴/۰۸/۱۷ تاریخ بازنگری: ۱۴۰۴/۱۰/۲۴ تاریخ پذیرش: ۱۴۰۴/۱۰/۲۸ تاریخ انتشار: ۱۴۰۵/۰۴/۲۷ کلیدواژه‌ها: فلسفه هوش مصنوعی، فقه و حقوق هوش مصنوعی، شمول‌گرایی، شهروند، افراد، اشخاص، فراگیری.	توسعه سریع هوش مصنوعی، توجهات را به شدت به پیامدهای فلسفی و حقوقی آن، به ویژه در مورد تبعیض و سوگیری، معطوف کرده است. یک ضعف اساسی در این گفتمان، تأکید بر مفهوم «شهروند» است که مقوله‌ای انحصاری بوده و تمامی ذی‌نفعان را در بر نمی‌گیرد. دستیابی به فراگیری واقعی در سیستم‌های هوش مصنوعی نیازمند تغییر بنیادین در اصطلاحات، از «شهروند» به مفهومی فراگیرتر است. عنوان «فرد»، جایگزین متداولی است، اما اصطلاح حقوقی «شخص» از فراگیری بالاتری برخوردار است. چون هم اشخاص حقیقی و هم اشخاص حقوقی را در بر می‌گیرد. این پژوهش با روش تحلیل مفهومی و انتقادی در چارچوب اندیشه فلسفی-اخلاقی، و فقهی-حقوقی و با ابزار کتابخانه‌ای تأکید می‌کند که محافظت از این فراگیری نه تنها یک ضرورت اخلاقی، بلکه یک الزام حقوقی است و به بررسی در مبانی فقه امامیه به عنوان نگرش حقوق شرعی می‌پردازد تا از رویکرد شخص‌محور پشتیبانی و آن را تبیین کند. این مقاله با تلفیق استدلال‌های فلسفی-اخلاقی، و فقهی-حقوقی، چارچوبی مستحکم‌تر برای طراحی و حکمرانی سیستم‌های هوش مصنوعی فراگیر ارائه داده است.

استناد: مومنی، محمدحسن و پرویزی، مریم. (۱۴۰۵). شمول‌گرایی در هوش مصنوعی از نگاه فلسفی - حقوقی، پژوهش‌های فلسفی، ۲۰ (۵۵)، ۲۲۱-۲۴۲.

<https://doi.org/10.22034/jpiut.2026.70131.4323>



© نویسندگان.

ناشر: دانشگاه تبریز.

مقدمه

رشد سریع هوش مصنوعی^۱ همراه با توجه ویژه به مباحث فلسفی - اخلاقی و حقوقی آن بوده است. انتشار روزافزون انواع راهنماهای اخلاقی و تلاش برای قانون‌گذاری در این حوزه خود نشان دهنده اهمیت این قبیل مسائل برای پژوهشگران است. با توجه به تأثیراتی که هوش مصنوعی بر جامعه می‌گذارد، یکی از مباحث بسیار مهم ذی‌نفعان است. در ادبیات پژوهشی چه در حوزه اخلاق و فلسفه اخلاق و چه در حوزه حقوق، ذی‌نفعان انسانی سیستم‌های هوش مصنوعی گاهی تحت عنوان شهروندان و گاهی تحت عنوان افراد موردبررسی قرار گرفته‌اند. از منظر حقوقی، شمول‌گرایی کارکرد و حقوق متفاوتی را برای ذی‌نفعان به ارمغان می‌آورد. بی‌تردید در نظام حقوقی ما مسئله رویکرد و مبانی فقهی یکی از دوال مرکزی است. پس باید شمول‌گرایی را علاوه بر منظر فلسفی - اخلاقی، از منظر فقهی - حقوقی مورد کنکاش قرار داد. این مقاله می‌کوشد تا شمول‌گرایی را به معنای دقیق کلمه که جنبه‌های فلسفی - اخلاقی و فقهی - حقوقی را حتماً در خود جای می‌دهد در سیستم‌های هوش مصنوعی معرفی کند. شایان‌ذکر است جنبه‌های فقهی مورد بحث ناظر بر فقه امامیه به‌عنوان یکی از دوال اصلی حکمرانی در ایران و از مبانی مهم نظام حقوق داخلی است. پس هر جا در این مقاله به جنبه فقهی - حقوقی توجه شد منظور فقه امامیه است.

پیشینه تحقیق

در سطح بین‌المللی پژوهش‌های متنوعی به مباحث فلسفی - اخلاقی هوش مصنوعی پرداخته است. اخلاق هوش مصنوعی (کوئکلیبرگ، ۲۰۲۰؛ مولر، الف ۲۰۲۱؛ فلوریدی، ۲۰۲۳)، اخلاق هوش مصنوعی در زمینه‌های اختصاصی (دیگنام، ۲۰۱۸؛ ژانگ و ژانگ، ۲۰۲۳؛ ژو و نابوس، ۲۰۲۳؛ نگوین، ۲۰۲۲)، و اخلاق هوش مصنوعی با رویکرد سیستم‌های اجتماعی فنی (سادک و همکاران، ۲۰۲۳؛ کوئکلیبرگ و ستر، ۲۰۲۳) از این شمار است. ادبیات پژوهشی حقوقی در حوزه هوش مصنوعی (پونکین و ردکینا، ۲۰۱۸؛ گرینشتاین، ۲۰۲۲) و همچنین سیستم‌های اجتماعی - فنی (باکینر، ۲۰۲۳؛ هولتون و بوید، ۲۰۲۱) نیز بسیار گسترده و قابل‌تأمل است. در زبان فارسی نیز، پژوهش‌های فراوانی مانند «پراگماتیسم، اخلاق هوش مصنوعی و سیستم‌های اجتماعی - فنی» (عاشوری کیسمی و پرویزی و مومنی، ۱۴۰۴)، مباحث عام اخلاق هوش مصنوعی (عاشوری کیسمی، ۱۴۰۳؛ عاشوری کیسمی و پرویزی، ۱۴۰۳؛ رضانی و فیضی درخشی، ۱۳۹۲) و اخلاق هوش مصنوعی در زمینه کاربردی (عباسی و تیموری، ۱۴۰۲؛ عاشوری کیسمی ۱۴۰۳ الف؛ سیفی و رزمخواه، ۱۴۰۲) صورت گرفته است.

در کتب فارسی نیز این خط سیر دنبال شده است مانند هوش مصنوعی: فلسفه، اخلاق و جامعه (صفدری شریانی، ۱۴۰۳). همچنین هوش مصنوعی از منظر حقوقی: تأثیر تحولات حقوقی هوش مصنوعی در حوزه حقوقی مالکیت فکری (اسدپور و همکاران، ۱۴۰۲) و هوش مصنوعی و صدور احکام کیفری: تصمیم‌سازی یا تصمیم‌گیری (شیخوند و همکاران، ۱۴۰۲). یکی از وجوه تمایز این مقاله با سایر پژوهش‌های صورت گرفته اولاً، توجه خاص به بحث شمول‌گرایی است. ثانیاً این بحث را در نگاهی جامع که متناسب هویت مقوله شمول‌گرایی است، هم از جهت فلسفی - اخلاقی و هم از جهت فقهی - حقوقی مورد مذاقه قرار داده است. جهت تمایز سوم آن است که شمول‌گرایی را بر اساس سیستم‌های اجتماعی - فنی نیز که بر پایه اقتضائات

^۱ Artificial Intelligence (AI): در مدخل اخلاق هوش مصنوعی در دانشنامه آنلاین استنفورد، هوش مصنوعی نوعی سیستم محاسباتی در نظر گرفته می‌شود که برای دستیابی به هدف/اهدافی طراحی شده و برای رسیدن به این اهداف، رفتار هوشمند از خود بروز می‌دهد در اینجا مقصود از هوشمند، به معنای سطح انسانی از هوشمندی نیست (مولر، الف ۲۰۲۱) ما نیز در این مقاله به سه علت از این تعریف تبعیت می‌کنیم اول آنکه تعریف ارائه شده، یکی از معتبرترین منابع مورد استفاده پژوهشگران است دوم تعریف میان حاضر هوش مصنوعی و سایر برنامه‌های رایانه‌ای را بر اساس رفتار هوشمند متمایز می‌کند و سوم اینکه تعریف حاضر جامعیت داشته و تعابیر هوش مصنوعی قوی و ضعیف را پوشش می‌دهد

ذات جامعه و مقتضیات نفس فن‌آوری متکی است روشن می‌کند و در این راستا، معرفی ذی‌نفعان را در قالب شهروند کافی و صحیح نمی‌داند، حتی معرفی آن‌ها را ذیل عنوان «فرد» ناقض هویت ذاتی شمول‌گرایی دانسته و به جهت لزوم شمول عنوان نسبت به هر دو وجه اشخاص حقیقی و حقوقی، عنوان «شخص و اشخاص» را پیشنهاد می‌کند.

هدف و ضرورت

هدف از پژوهش حاضر بررسی شمول‌گرایی فلسفی- اخلاقی و فقهی- حقوقی در حوزه هوش مصنوعی است که در این راستا توجه به رویکرد سیستم‌های اجتماعی- فنی نیز مد نظر بوده است. در ادبیات پژوهشی، اگرچه شمول‌گرایی بارها توصیه شده است، اما فرضیه این تحقیق بر آن است که چه از منظر فلسفی- اخلاقی و چه در نگاه فقهی- حقوقی و بخصوص بر اساس رویکرد سیستم‌های اجتماعی- فنی، لازمه عملکرد اخلاقی و حقوقی سیستم‌های هوش مصنوعی عدم تمرکز بر ذی‌نفعان شهروند و توجه بر تمامی ذی‌نفعان صرف‌نظر از وضعیت شهروندی آن‌ها است. در راستای رسیدن به هدف اصلی، مقاله حاضر در سه بخش تدوین شده است. بخش اول، رویکرد سیستم‌های اجتماعی فنی و لازمه شمول‌گرایی. بخش دوم، شمول‌گرایی در سیستم‌های اجتماعی فنی به عنوان یک رویکرد «فلسفه پایه» در ارتباط با مباحث اخلاق هوش مصنوعی. بخش سوم، شمول‌گرایی در حوزه هوش مصنوعی از منظر فقهی- حقوقی، با نظر داشت رویکرد سیستم‌های اجتماعی فنی. ضرورت پژوهش حاضر از آن روست که تاکنون پژوهش مستقلی به لزوم شمول‌گرایی در هوش مصنوعی، هم از منظر فلسفی- اخلاقی و هم از منظر فقهی- حقوقی، بخصوص با توجه به سیستم اجتماعی- فنی نپرداخته است.

۱. سیستم‌های اجتماعی فنی و شمول‌گرایی

۱-۱. سیستم‌های اجتماعی فنی

مفهوم سیستم اجتماعی فنی در ادبیات مختلفی از جمله فناوری در محیط کار، مهندسی سیستم‌ها، مطالعات علم و فناوری، فلسفه فناوری، اخلاق مهندسی و پژوهش‌های هوش مصنوعی و سیستم‌های خودمختار به کار رفته است. آنچه در این مطالعات مشترک است این است که علی‌رغم اختلاف در نحوه توجه و تغییر مفهوم سیستم‌های اجتماعی فنی، همگی به این اذعان دارند که علاوه بر عوامل فناورانه، عوامل اجتماعی مختلفی نیز در حوزه فن‌آوری، (بخصوص در فن‌آوری‌های دیجیتال و به‌عنوان فن‌آوری شاخص، در سیستم‌های هوش مصنوعی)، مورد توجه است (کودینا و ون دو پول، ۲۰۲۴، ۲).

مفهوم سیستم‌های اجتماعی فنی بر ۲ اصل کلی استوار است: ۱- نظام‌مندی فناوری‌ها، به این معنا که فناوری‌های مختلف، بخشی از سیستم‌های بزرگتر (فنی یا اجتماعی- فنی) است. هر سیستم نیز از مجموعه‌ای از عناصر (فناوری‌ها، انسان‌ها، قوانین) شکل گرفته است که با یکدیگر تعامل دارند و هدف مشترکی را دنبال می‌کنند. برای مثال در سیستم انرژی نیروگاه، مصرف کننده و ...، در سیستم حمل‌ونقل، کاربران و ...، در سیستم‌های بانکی از بانک جهانی گرفته تا یک شعبه از یک بانک خاص. ۲- محوریّت عناصر اجتماعی؛ هر سیستم اجتماعی- فنی عملکردش به عامل‌های انسانی/ اجتماعی وابسته است. وجه مشترک همه رویکردهای مختلف در سیستم‌های اجتماعی فنی، همین ارتباط پیچیده میان مؤلفه‌های اجتماعی و فنی در کنار یکدیگر است سیستم‌های هوش مصنوعی به‌عنوان یک سیستم اجتماعی فنی، به دلیل ماهیت متفاوت هوش مصنوعی، مؤلفه‌های بنیادین اضافه‌تری دارد. سه ویژگی خودمختاری^۱، تعاملی بودن^۲ و انطباق‌پذیری^۳ تفاوت اصلی هوش مصنوعی با فناوری‌های سنتی است.

¹ autonomous

² interactive

³ adaptive

باید توجه داشت که هوش مصنوعی مانند سیستم‌های اجتماعی فنی سنتی توانایی یادگیری از محیط و تطبیق با آن را دارد. اما تفاوت در این است که در سیستم‌های اجتماعی فنی سنتی این امر با عامل‌های انسانی محقق می‌شود، اما سیستم‌های هوش مصنوعی دارای عناصر فنی است که امکان یادگیری مستقیم از محیط را فراهم می‌کند. دو مؤلفه سازنده اضافی در سیستم‌های هوش مصنوعی، عامل‌های مصنوعی^۱ و قواعد فنی^۲ است. عامل‌های مصنوعی همان عوامل غیرانسانی^۳ و غیرعاملی^۴ هستند که می‌توانند خودمختار عمل کنند. و در مقابل ورودی‌های محیط خود را منطبق کنند. قواعد فنی، بین عامل‌های مصنوعی و دیگر فناوری‌ها ایجاد تعامل کرده و به تنظیم آن اقدام می‌کند. می‌توان نقش این قواعد در سیستم‌های هوش مصنوعی را مشابه نقش نهادهای اجتماعی در سیستم‌های اجتماعی فنی سنتی دانست (کودینا و ون دو پول، ۲۰۲۴، ۳).

نکته قابل توجه این است که با این تحلیل می‌توانیم ماهیت منحصر به فرد سیستم‌های هوش مصنوعی (یادگیری و انطباق نه فقط به وسیله انسان‌ها بلکه از طریق مکانیسم‌های فنی خود سیستم) را درک کنیم.

۱-۲. رویکرد سیستم‌های اجتماعی فنی و هوش مصنوعی

«اولیا کودینا»^۵ و «ایبو فان دِ پول»^۶ معتقدند درک عمیق‌تر (از نحوه عملکرد تا تأثیرات اجتماعی، چالش‌های اخلاقی و راه‌حل‌های احتمالی) اولین و مهم‌ترین مزیت رویکرد سیستم اجتماعی-فنی در حوزه فن‌آوری هوش مصنوعی است. رویکرد سیستم اجتماعی فنی بر این نکته تأکید دارد که هوش مصنوعی در بستر گسترده‌تری از سیستم‌های اجتماعی فنی قرار دارد و به طراحی بهتر آن کمک می‌کند. اما اهمیت این رویکرد فراتر از درک عمیق است. زیرا دگرگونی‌هایی که هوش مصنوعی در این سیستم‌ها (مثلاً انرژی یا حمل‌ونقل و ...) ایجاد می‌کند، جامعه و نهادهای اساسی همچون قانون‌گذاری و دموکراسی را نیز تحت تأثیر قرار می‌دهد. با فرض گفته شده در مورد اینکه هوش مصنوعی مؤلفه‌های بنیادین جدیدی به سیستم‌های اجتماعی فنی اضافه می‌کند، باید منتظر تغییرات اساسی‌تری از نوآوری در فناوری و تکنولوژی از طرف آن باشیم (کودینا و ون دو پول، ۲۰۲۴، ۲).

۱-۲-۱. خطرات

هوش مصنوعی، فناوری نوآورانه و تحول‌آفرینی است که می‌تواند سیستم‌های اجتماعی-فنی دیگر و یا حتی ساختار جامعه را دستخوش تغییر کند (هاپستر، ۲۰۲۱، ون دو پول و همکاران، ۲۰۲۳).

برای مثال می‌توان از مختل کردن شیوه‌های موجود نوشتن و آموزش توسط چت جی پی تی (شیدک، ۲۰۲۳) یا به چالش کشیدن تمایز بین موجودات انسانی و غیرانسانی مثل ربات‌ها، توسط ربات‌های اجتماعی (نیپولم، ۲۰۲۳) نام برد. خطرات هوش مصنوعی به مبحثی جدی و مستقل تبدیل شده و حقوقدانان و فیلسوفان بر این مقوله به‌طور جدی عطف توجه کرده‌اند، مانند «نیک بوستروم»^۷ (بوستروم، ۱۳۹۱، ۸۳ و ۸۴).

¹ artificial agents

² technical rules

³ non-human

⁴ non-intentional

^۵ Kudina, O استاد دانشگاه صنعتی دلفت و پژوهشگر در حوزه اخلاق و فلسفه تکنولوژی

^۶ Van de Poel, I استاد دانشگاه صنعتی دلفت و پژوهشگر در حوزه اخلاق و فلسفه مهندسی و مسئولیت اجتماعی تکنولوژی

⁷ Nick Bostrom

۱-۲-۲. لزوم توجه به رویکرد اجتماعی فنی در هوش مصنوعی

رویکرد سیستم اجتماعی فنی سه عنصر را به طور جدی مورد توجه قرار می‌دهد: نهادها یا مؤسسات^۱، فرهنگ^۲ و حکومت^۳. نهادهای اجتماعی یا مؤسسات به این دلیل که در تنظیم تعاملات انسانی و فناوری نقش حیاتی دارد مهم است. فرهنگ به عنوان یک نهاد فراتر از نهادهای غیررسمی دیگر به توجه ویژه نیاز دارد. برخورد صحیح با مسائل اخلاقی و اجتماعی مطروح توسط هوش مصنوعی صرفاً به ساز و کارهای اخلاقی و طراحی مناسب مربوط نمی‌شود، بلکه باید به مسائل حکومتی و سیاست‌گذاری (حکمرانی) توجه ویژه‌ای داشت. در مقابله با استعدادهای مخرب هوش مصنوعی به ساز و کاری بیشتر از اخلاق نیاز داریم که توجه ویژه‌ای به طیف وسیعی از عوامل اجتماعی، اقتصادی، حاکمیتی و... داشته باشد. این مهم میسر نیست مگر آنکه نخست، توجهی بنیادین و فلسفی به حوزه اخلاق در کنار توجه کاربردی به آن در حوزه هوش مصنوعی تحقق یابد. ثانیاً، رویکرد حقوقی به حوزه هوش مصنوعی مورد توجه جهانی به صورت یکپارچه، نه صرفاً پرداختی جزیره‌ای و زیست‌بوم‌گرایانه واقع شود. البته همخوانی با زیست‌بوم‌های فرهنگی امری گریز ناپذیر است لذا توجه به مبانی و قواعد فقهی در حوزه حقوق داخلی بسیار حائز اهمیت است. چرا که این نهادها خود از مؤلفه‌های اجتماعی موجود است که بخصوص در نگاه سیستم‌های اجتماعی-فنی از ارکان مهم اثرگذار در پرداخت به حوزه فن‌آوری، به‌ویژه فن‌آوری‌های نوین هوش مصنوعی است. نقش بنیادهای فقهی، (به‌طور خاص فقه امامیه)، اعم از مبانی و قواعد ماهوی آن در جریان نظام حقوقی حاکم، درواقع بخش مهمی از نظام فرهنگی-اجتماعی را شکل می‌دهد. بنابراین در ایران بخصوص با رویکرد اجتماعی-فنی نمی‌توان از نقش فقه و حقوق در تنظیم روابط موجود بین انسان و هوش مصنوعی غفلت کرد، کما اینکه چنین تنظیمی حتماً به مبانی فکری فلسفی-اخلاقی نیازمند است. این مهم در ساختار سیستم اجتماعی-فنی بیشتر خود را نشان می‌دهد زیرا لایه‌های تفکر فلسفی در نگرش اخلاقی جامعه به فن‌آوری‌های نوین را نمی‌توان از بافت اجتماعی-فرهنگی جامعه جدا و دور کرد.

مدیریت هوش مصنوعی به توسعه سیاست‌های هوشمند یکدست نیازمند است (کودینا، ون دو پول، ۲۰۲۴، ۴). پس نگاه طراحی و توسعه به هوش مصنوعی باید متوجه بسترهای اجتماعی فنی که در آن تعبیه شده است باشد. اینجاست که شمول‌گرایی به عنوان عاملی اساسی در مقابل سوگیری و در جهت یکدست شدن سیاست‌های مورد نیاز باید مورد توجه قرار گیرد.

۱-۳. شمول‌گرایی

شمول‌گرایی، به عنوان عاملی در برابر سوگیری می‌تواند عمل کند. هوش مصنوعی نباید محدود به اطلاعات خاص طبقه‌بندی شده و مقید به جهت‌گیری خاصی باشد. مثلاً نژاد، رنگ و ملیت افراد در برنامه‌هایی که با ابزار هوش مصنوعی انجام و مدیریت می‌شود مؤثر نباشد.

برای مثال، گاه طبق الگوریتم‌های پزشکی، سوگیری نژادی در مراقبت‌های بهداشتی اتفاق می‌افتد و بیماران سفیدپوست نسبت به بیماران سیاه‌پوست ارجحیت پیدا می‌کنند (جانسون، ۲۰۱۹) یا سوگیری تبعیض جنسیتی در توصیه‌نامه‌ها در یک آزمایش استخدام توسط چت جی پی تی کاملاً مشهود است (استوکل-واکر، ۲۰۲۳).

اگر داده‌های هوش مصنوعی متنوع و نماینده تمامی افراد (به دور از جنسیت، نژاد، توانایی و...) نباشد و هوش مصنوعی از ظرفیت شمول فراگیر عاری باشد، به نحو سیستماتیک دچار تبعیض می‌شود.

به‌طور کلی می‌توان گفت شمول‌گرایی و تنوع در هوش مصنوعی را می‌توان به ۴ دلیل ضروری دانست:

¹ institutions

² Culture

³ governance

➤ **کاهش سوگیری‌ها:** سیستم‌های هوش مصنوعی که دارای داده‌های ناقص یا جانبدارانه هستند باعث تبعیض می‌شوند و این در تصمیم‌گیری‌های مهم اجتماعی مثل استخدام یا موارد حقوقی و قضایی بسیار مهم و در جهت منفی و سوگیری اثرگذار است. در نظام فقه امامیه هرگونه سوگیری به معنای در نظر داشت حق برتری و تقدم بی‌وجه برای هر مورد و موضوعی مردود و از شمار منکرات است. سوگیری، به معنای ایجاد امتیاز و تمایز بخصوص در حقوق اساسی و انسانی به دلیل نقض عدالت قابل‌پذیرش در فقه امامیه نیست. آموزه‌هایی نظیر نفی برتری نژادی و رنگی و قومی و جنسیتی از شواهد گویای این مهم در منابع معتبر فقه امامیه است (حجرات / ۴۹: ۱۳، قرطبی، ۱۶، ۳۴۲). بند نهم و چهاردهم از اصل سوم قانون اساسی بر رفع هرگونه تبعیضی صراحت دارد. این اصل چون در مقام بیان قانونی است، پس دارای جنبه اطلاقی بوده و شامل تبعیض‌ها و سوگیری‌های ممکن در فن‌آوری‌های نوین از قبیل سیستم‌های هوش مصنوعی می‌گردد و نمی‌توان به دلیل تقدم تصویب این اصل بر حضور این دست فن‌آوری‌ها آن را فاقد وجهت قانونی دانست. این اصل در چهره دیگری در قالب اصل نوزدهم و بیستم قانون اساسی نیز بروز یافته و همه آحاد جامعه را از جهت حقوق انسانی، سیاسی، اقتصادی، اجتماعی و فرهنگی از حمایت مساوی قانونی برخوردار می‌داند.

➤ **تقویت انصاف:** شمول‌گرایی در هوش مصنوعی ضامن بهره‌مندی تمامی اقشار جامعه از آن است، و تمامی افراد از تکنولوژی به‌طور برابر بهره‌مند می‌شوند و از این جهت انصاف همگانی بیشتر رعایت شده است. بنابر آنچه متخصصان یونسکو تصریح کرده‌اند هوش مصنوعی در حرکت به سوی فراگیری در تمام شئون زندگی بشری است (گروه متخصصان یونسکو، ۱۴۰۱، ۱۸). رعایت و تقویت انصاف در چنین فضا و شرایطی جدای از آنکه با روح فلسفه اخلاق بسیار سازگار و نزدیک است و ضرورت‌های اخلاقی را تأمین می‌کند، از جهت فقهی - حقوقی نیز در جهت رعایت مسئولیت‌های اجتماعی و فردی اثرگذار است و از سنخ تضمین نرم برای پیشگیری از بسیاری از جرائم و تخلفات قانونی می‌تواند باشد. به اعتقاد برخی نقش انصاف را باید در تضمین عدالت جستجو کرد (امینی، ۱۳۹۵، ۳۳). تضمین عدالت از رهگذر انصاف، در سطح جامعه به شکل‌دهی مؤلفه ضروری در سیستم اجتماعی - فنی درباره هوش مصنوعی یاری می‌رساند و فلسفه شمول‌گرایی در فن‌آوری را مبرهن می‌سازد.

➤ **افزایش خلاقیت:** اندیشه پنهان درباره افزایش خلاقیت و رابطه آن با شمول‌گرایی بر پایه این فلسفه استوار است که با تنوع و افزایش گروه‌های توسعه دهنده هوش مصنوعی و تنوع دیدگاه‌های مختلف می‌توان به نوآوری و خلاقیت دست یافت و طیف وسیعی از کاربران را پوشش می‌دهد. اگر به مقوله هوش مصنوعی از منظر یک موهبت، (موهبتی تکنولوژیک)، نگاه کنیم جای این معنای حقوقی فراهم می‌آید که در بند دوم از اصل سوم قانون اساسی از جهت بالا بردن سطح آگاهی‌های عمومی در عرصه رسانه و سایر وسایل (از جمله هوش مصنوعی) و در بند چهارم از همین اصل در جهت تقویت روح ابتکار در تمام زمینه‌های علمی و فنی مورد تأکید قرار گرفته است. از منظر فقهی توجه به حدیث معروف «الناس معادن كمعادن الذهب و الفضة» (ثقة الاسلام الکلینی، ۱۳۶۳، ۸، ۱۷۷) مقتضی آن است که در راستای بهره برداری از این معادن ذیل وجوب عقلی مقدمی به مقوله خلاقیت و افزایش آن اهتمام وافی شود و از آنجائی که شمول‌گرایی چنین مقصودی را فراهم می‌سازد طبیعی است که از حمایت مبنای فقهی در راستای تفسیر روایت مذکور برخوردار باشد (مکارم، ۱۳۸۵، ۲، ۳۳۱).

➤ **پیشگیری از حوادث اجتماعی زیان‌بار:** عدم فراگیری و شمول سیستم‌های هوش مصنوعی باعث به عقب رانده شدن اقلیت‌ها و طیف گروه‌های در حاشیه و کم تعداد شده و این خود سبب آسیب به آن‌ها می‌شود و این یعنی نقض

حقوق بشر و افزایش نابرابری‌های اجتماعی (زوقی و بانو، ۲۰۲۴، ۱). مسئله به حاشیه راندن اقلیت‌ها و نادیده گرفته شدن حقوق آن‌ها در نگاهی کلان به آسیب کل جامعه منتهی می‌شود. در فلسفه موسوم به فلسفه اسلامی، تحقق کمال انسانیت در گرو توجه به حقیقت وجود مطلق است. وجود مطلق حقیقت واحد مشککی است که همه ماهیات متکثره را در برمی‌گیرد (طباطبایی، بی تا، ۱۵). بنابراین شمول‌گرایی از این جهت کاملاً همداستان با نظام فلسفی قابل تعریف است. از جهت فقهی - حقوقی نیز به حکم قاعده لا ضرر و نفی عسر و حرج هرگونه ایجاد حرج و سختی طاقت‌فرسا و هرگونه ایجاد ضرر بر هر فرد یا گروهی مورد نهی است و از نگاه فقهی تحمل ضرر و حرج چه از جهت عقلی و چه شرعی پذیرفته نیست (نجفی، ۱۴۲۱ ق، ۲۰، ۲۴۵).

نکته حائز اهمیت در بحث شمول‌گرایی نسبت به مقوله سوگیری و تبعیض و هر امر دیگری در سیستم‌های اجتماعی فنی آن است که این دست عوامل در جایگاه اثر و فایده باید دیده شود؛ یعنی چنین نیست که ارزش شمول‌گرایی در گرو این قبیل عوامل تعریف گردد. خیر، بلکه شمول‌گرایی در سیستم اجتماعی - فنی به حکم اقتضای ذاتی این سیستم چون امری بر آمده از بافت اجتماعی و نیاز تعامل همه‌جانبه با فن‌آوری هوش مصنوعی است دارای بار ارزشی است و مسئله ارزش باری^۱ در این مقوله در گروی صرفاً نفی سوگیری و نظایر آن نیست.

۲. اخلاق فلسفی، مبنای فقهی - حقوق و شمول‌گرایی در هوش مصنوعی

۲-۱. نظام اخلاقی - حقوقی در هوش مصنوعی

«مولر» موضوعاتی چون حریم خصوصی و نظارت^۲، سوگیری در تصمیمات سیستم^۳، تاری^۴ و شفافیت^۵، دستکاری در رفتار^۶، سیستم‌های خودمختار^۷، اتوماسیون و اشتغال^۸ و تکینگی^۹ را از موارد پرتکرار در مورد پژوهش‌های حوزه فلسفه و اخلاق هوش مصنوعی می‌داند (عاشوری کیسی، ۱۴۰۳، ۹۸ و ۹۹؛ مولر، ۲۰۲۱ ب). موارد مذکور هر یک افزون بر جنبه‌های اخلاقی و فلسفی، نسبت به جنبه‌های فقهی - حقوقی نیز اثرگذار است.

۲-۱-۱. حریم خصوصی و نظارت در رویکردی فلسفی - اخلاقی، و فقهی - حقوقی

بحث از حریم خصوصی و نظارت بیشتر حول محور دسترسی به اطلاعات و داده‌های خصوصی است که با گسترش تکنولوژی و جمع‌آوری و ذخیره داده‌ها به صورت غیر دستی و دیجیتال، تجزیه و تحلیل این اطلاعات سهل‌تر شده و بسیار افزایش یافته است. اطلاعات خصوصی افراد در بانک‌های داده شرکت‌ها ذخیره می‌شود و با استفاده کاربران از وسایل دیجیتال و هوش مصنوعی؛ داده‌ها طبقه بندی شده و گاه برخی شرکت‌ها این داده‌ها را با هم داد و ستد می‌کنند به نحوی که علم به اینکه اکنون چه کسانی به این داده‌ها دسترسی دارند دشوار شده است (مولر، ۲۰۲۱ الف). با یک جستجوی ساده در اینترنت در مورد خرید لوازم خانگی یا ملزومات دیگر، داده‌های طبقه بندی شده به شرکت‌های فروش این کالاها ارسال شده و پس از جستجوی فرد، وی شاهد دریافت

^۱ Value Alignment

^۲ Privacy and surveillance

^۳ Bias in decision systems

^۴ Opacity

^۵ Transparency

^۶ Manipulation of behavior

^۷ Autonomous systems

^۸ Automation and employment

^۹ Singularity

پیامک‌های تبلیغاتی این قبیل کالاها خواهد بود. این امر باعث نگرانی افراد از ذخیره و تحلیل داده‌های خصوصی آن‌ها شده و به یکی از چالش‌های مهم در هوش مصنوعی تبدیل شده است.

در فقه امامیه هرگونه تجسس مورد نهی است (نجفی، ۱۴۲۱ق، ۷، ۲۲۵) و در قرآن کریم از آن نهی کرده است.^۱ تجسس در امور دیگران، ناقض حریم خصوصی است. در حمایت از حریم خصوصی قاعده فقهی سلطنت نیز جاری و حاکم است (شیخ انصاری، ۱۴۱۵ق: ج ۳، ۴۱ و ۸۷ و ۹۸، و ج ۵، ۲۰، و ج ۶: ۲۱۶). مطابق این قاعده هر فردی نسبت به اموال و انفس خود دارای سلطنت است. نقض حریم خصوصی نادیده گرفتن و شکستن چنین سلطنتی است. نفی شمول‌گرایی می‌تواند در جهت نقض حریم خصوصی و شکستن سلطه و سلطنت پذیرفته شده اشخاص عمل کند و از این جهت شمول‌گرایی مورد تأیید و تقویت نظام فقهی است. قاعده فقهی سلطنت علیرغم عدم احراز خاستگاه روایی‌اش، چه در منابع روایی امامیه و چه در منابع روایی اهل تسنن، حکم و مفهومی عقلانی است (بروجردی، ۱۴۱۳ ق، ۱۸۰) و از این جهت رابطه مستقیم عرفی - اجتماعی و عقلانی بین حفظ حریم خصوصی و ویژگی شمول‌گرایی آشکار می‌گردد.

مسئله شمول‌گرایی در وابستگی به عنصر حفظ حریم خصوصی نمی‌تواند بر مبنای سیستم‌های اجتماعی - فنی تکیه کند. زیرا در این سیستم ذات شمول‌گرایی فارغ از جنبه حریم خصوصی، به دلیل اقتضای فرهنگی و اجتماعی ذاتی‌اش دارای بار ارزشی است. به‌هرحال برای شمول‌گرایی متکی بر عنصر پرهیز از نقض حریم خصوصی می‌توان ریشه فقهی دیگری را نیز تصویر کرد. قاعده لزوم ستر عورت (نجفی، ۱۴۲۱، ۷، ۲۲۵) یا به اطلاق لفظی و یا به وحدت ملاک قابل استناد در این خصوص است. اگر اطلاق عورت را شامل هر امری بدانیم که برای اشخاص وجاهت خصوصیت دارد می‌تواند شامل مطلق حریم خصوصی باشد و ستر و پوشیده ماندنش تکلیفی بر همگان است که نباید به این حریم تعدی و تجاوز کنند. نادیده گرفتن شمول‌گرایی به نوعی به نقض چنین حریمی می‌انجامد و از باب تحریم عقلی و شرعی ارشادی مورد نهی و منع است. حمایت از شمول‌گرایی در سیستم‌های اجتماعی فنی نیز فاقد وجاهت فقهی نیست. چراکه عموم جامعه و منطق اجتماعی پیوند حفظ حریم خصوصی با شمول‌گرایی را واجد ملازمه عرفی می‌داند. در ابواب غیرعبادی، عرف و اعتبارات عرفی اگر مورد نهی و ردع شرعی قرار نگرفته باشد حائز اهمیت است و عرف عمومی در چنین موقعیتی واجد اعتبار فقهی است (حسینی، بی‌تا، ۴، ۵۱ و ۲۲۹. میرداماد، ۱۳۹۳، ۱، ۲۴۶. سبحانی، ۱۳۹۷، ۱۶۷ تا ۱۷۲). فقط در جلد اول قانون مدنی بر اعتبار و رجوع به «عرف و عادت و متعارف» در مواد و مواضع گوناگون، سی‌ویک بار تصریح شده و از خصوص «عرف» بیست‌وهفت بار در مواد گوناگون یاد شده است.^۲ در خصوص نقض حریم خصوصی، در ماده ۱۰۷ آئین دادرسی کیفری هرگونه تعقیب و تحقیق در جرائم منافی عفت ممنوع اعلام شده و پرسش در این خصوص از هیچ فردی مجاز نیست. بنابر چنین بن مایه فقهی و حقوقی می‌توان سیستم‌های اجتماعی - فنی را در خصوص شمول‌گرایی تأیید کرد. وقتی مقتضای فرهنگ عمومی در تعامل با فناوری و اقتضائات آن اصرار بر شمول‌گرایی باشد چنین معنایی واجد اعتبار عرفی شده و مورد نهی و ردع شرعی نیز نیست. از سوی دیگر چون در قلمرو موضوعات غیرشرعی ورود دارد و به وضع و جعل حکم شرعی نمی‌پردازد نمی‌توان عرف را در حصار عرف عهد شارع مقدس منحصر کرد و از این باب وجاهت عرفی شمول‌گرایی را عاری از حمایت فقهی تصور کرد. از آنجا که سوگیری^۳ یکی از نقاط مقابل شمول‌گرایی است ضروری است به طور خاص به بررسی انواع سوگیری در هوش مصنوعی پرداخته و از ابعاد فلسفی اخلاقی و فقهی حقوقی مورد بررسی قرار دهیم.

^۱ حجرات/۴۹: ۱۲

^۲ مواد ۱۳، ۵۴، ۱۶۴، ۲۲۰، ۲۲۴، ۲۲۵، ۲۸۰، ۳۳۲، ۳۳۶، ۳۴۲، ۳۴۴، ۳۵۶، ۳۵۷، ۳۵۸، ۳۵۹، ۳۶۹، ۳۷۵، ۳۸۲، ۴۱۷، ۴۲۶، ۴۸۶، ۵۲۴، ۵۴۲، ۵۴۹، ۵۵۵، ۵۶۱، ۶۲۳، ۶۶۶، ۶۶۷، ۶۷۶ و ۹۵۰

^۳ Bias

۲-۱-۲. سوگیری در تصمیمات سیستم^۱ در رویکردی فلسفی- اخلاقی، و فقهی- حقوقی

«مولر»^۲ نوع سوگیری مؤثر در هوش مصنوعی را در جریان یادگیری و خروجی ماشین از هم تمیز می‌دهد: سوگیری آموخته‌شده^۳، سوگیری شناختی^۴ و سوگیری آماری^۵.

از سوگیری‌های تبعیض‌آمیز در مورد اقشار و افراد خاصی از جامعه و قضاوت ناعادلانه تحت تأثیر موضوعات نامرتبط به سوگیری آموخته‌شده یاد می‌شود. در این نوع سوگیری که اغلب ناخودآگاه صورت می‌گیرد ممکن است سوگیری در فرایند طراحی، گزینش داده‌ها و یادگیری ماشین اتفاق بیفتد. از آنجایی که نوع بشر در معرض بسیاری از سوگیری‌های شناختی است، چنین سوگیری رخ می‌دهد. افراد مایلند داده‌ها را در جهت تأیید به آنچه پیش‌تر باور و اعتقاد داشته و پیش‌فرض آنان بوده است تفسیر و تحلیل کنند (مولر، ۲۰۲۱الف).

در مورد سوگیری آماری می‌توان گفت سوگیری در طراحی ماشین نیست بلکه به دلیل موارد کثیری از یک داده که به سمت یک مورد خاص سوق داده شده‌اند اتفاق می‌افتد. مثلاً اینکه ماشین در مورد ساکنان آفریقا تماماً خروجی سیاهپوست بدهد زیرا عمدتاً داده‌های تصویری از آفریقایی‌ها سیاهپوست بوده است. یا در صورتیکه چهره یک فرد سیاسی یا مذهبی خاصی که در گذشته‌های دور می‌زیسته‌اند از ماشین خواسته شود، تصویر خروجی به دلیل کثرت داده‌ها زیبا یا زشت باشد حتی اگر چهره واقعی فرد هرگز دیده نشده باشد.

از منظر فقهی- حقوقی مسئله نفی سوگیری قابل استناد به قواعد گوناگون فقهی است. اصل کرامت انسانی که در قرآن کریم^۶ مورد^۷ تصریح واقع شده کمترین سوگیری را از جهت نقض عدالت و ایجاد ستم بر هر شخصی در شمار منکرات قرار می‌دهد. قاعده حرمت اعانه بر اثم نیز سوگیری را به عنوان ظلمی فاحش نسبت به افراد متضرر از آن ممنوع اعلام می‌کند (مأذبه: ۲، خوئی، ۱۴۱۸ق، ۳۵، ۲۶۸. بجنوردی، ۱۳۷۷، ۱، ۳۵۹). دلالت قاعده عدل و انصاف یا احسان، مستند به قرآن کریم^۸ نیز بر این مهم روشن است (هاشمی، ۱۳۸۲، ۶، ۲۶۶). سوگیری نیز مانند نقض حریم خصوصی موجب ضرر و زیان اشخاص و گروه‌هایی می‌گردد و مشمول نهی و نفی قاعده لاضرر است. تمام قواعد مذکور افزون بر جایگاه فقهی دارای وجاهت حقوقی، بخصوص در سطح اصول قانون اساسی و فراتر از مواد قوانین موضوعه است. بند ج از اصل دوم و بند ششم و نهم و چهاردهم از اصل سوم، اصل نوزدهم و بیستم و چهلم همگی بر این معنا و مقصود دلالت واضح دارد.

دغدغه جدی در سیستم‌های اجتماعی- فنی در مطلوبیت ذاتی عنصر شمولگرایی در نهاد جامعه و اقتضای تعامل آن با حوزه فناوری‌های نوین است. این منظور نیز در قلمرو فقه و نظام حقوقی دارای بن مایه و ریشه است. یکی از این بن مایه‌ها قاعده حسن و قبح عقلی است (مظفر، ۱۳۸۷، ۲۱۸ و ۲۲۴ و ۲۲۵). بی‌تردید سوگیری و همچنین تبعیض بر اساس هر عاملی در فرهنگ عرف عام و از نگاه عقلی بر اساس این قاعده محکوم و مردود است. لحاظ قاعده حسن و قبح عقلی در ردیف قضایای مشهوره و آرای محموده، آن‌سان که برخی اصولیین مانند مظفر (مظفر، ۱۳۸۷، ۲۳۴) و محقق اصفهانی (محقق اصفهانی، ۱۴۲۹ق، ۳، ۱۸)، و بسیاری از فلاسفه مانند بوعلی سینا (ابن‌سینا، ۱۳۸۶، ۱، ۲۱۹ و ۲۲۰) و علامه طباطبائی (طباطبائی، ۱۳۶۸، ۸: ۵۴) بر این باورند، این محکومیت را در لایه‌های عرفی جستجو می‌کند و از خاستگاه عقلی جامعه می‌نگرد. با توجه به چنین

^۱ باید توجه داشت که ممکن است بسیاری از سوگیری‌های مطرح با هم همپوشانی یا دارای مصداق‌های مشترک باشند

^۲ Learned Bias

^۳ Cognitive Bias

^۴ Statistical Bias

^۵ اسراء: ۷۰ «وَلَقَدْ كَرَّمْنَا بَنِي آدَمَ»

^۶ نحل: ۹۰ «إِنَّ اللَّهَ يَأْمُرُ بِالْعَدْلِ وَالْإِحْسَانِ»

مبنایی در علم اصول و فقه و دانش فلسفه می‌توان این قاعده را در واقع تأمین‌گر دغدغه سیستم‌های اجتماعی - فنی در خصوص عنصر شمول‌گرایی دانست.

۲-۱-۳. تاری (کدري) و شفافیت در رویکردی فلسفی - اخلاقی، و فقهی - حقوقی

در راستای توجه به اخلاق داده‌ها^۱ و اخلاق کلان داده‌ها^۲ مبحث تاری یا کدري و شفافیت در اخلاق هوش مصنوعی به میان می‌آید. اگر سیر و روند رسیدن ماشین به یک نتیجه خاص مشخص باشد، «شفافیت» و اگر این روند نامشخص باشد، «تاری و کدري» است (مولر، ۲۰۲۱الف). در یادگیری ماشین^۳ سه نوع کدري و تاری قابل تمییز از یکدیگر است: کم‌عمق^۴، استاندارد^۵ و عمیق^۶ (مولر، ۲۰۲۱ب، ۱۸). عدم شفافیت می‌تواند زمینه‌ای برای انواع غبن و غرر و فریب و نیرنگ و خدعه و ضرر و زیان گردد و از آنجائی که نظام فقهی - حقوقی در پی آثاری از این دست مواجهه و برخورد قانونی می‌کند، طبیعی است که چنین امری را بر نمی‌تابد. لزوم علم و آگاهی در کلیه روابط حقوقی قراردادی و هر امر دایر بر رضا و اراده نیز مقتضی لزوم شفافیت موجب علم و آگاهی است. این معنا در بند سوم ماده ۱۹۰ و ماده ۲۱۶ قانون مدنی قابل استنتاج است. ماده ۲۹۶ قانون آئین دادرسی مدنی در خصوص رأی دادگاه رعایت پنج بند و امر را ضروری دانسته که همگی در جهت شفافیت رأی مؤثر است. ماده ۴۰۳ همین قانون به‌نوعی بر اشراف دیوان عالی بر آرای دادگاه‌ها نظر دارد. روشن است که احاطه بر این معنا و فرایند زمانی قابل تصور است که آرای دادگاه‌ها دارای شفافیت لازم باشد تا دیوان بتواند به تأیید آنها اقدام کند. ماده ۵ قانون آئین دادرسی کیفری درباره لزوم بهره‌مندی متهم از حقوق دفاعی خویش درواقع به ملاک شفافیت در موضوع اتهام و ادله اتهام انتسابی اشاره دارد. کما اینکه ماده ۳۷۴ همین قانون در خصوص رأی دادگاه بر لزوم همین ملاک تکیه دارد.

از منظر فلسفی نیز بخصوص در مکتب صدرایی گوهر علم و آگاهی از صفات کمال و امتیازات وجودی و عین کمال نفس در مرتبه وجودی آن علم است (ملاصدرا، ۱۳۶۸، ۲، ۲۳۵). بنابراین، تفکر فلسفی به ذات شفافیت از جهت عنصر علم و آگاهی اهمیت داده و در حوزه هوش مصنوعی می‌توان این ویژگی را واجد اهمیت فلسفی مذکور دانست.

۲-۱-۴. دستکاری در رفتار در رویکردی فلسفی - اخلاقی، و فقهی - حقوقی

دستکاری در رفتار موقعی اتفاق می‌افتد که علیرغم وجود داده‌های کافی، سوگیری رفتاری، ایجاد اعتیاد و حتی فریب تبلیغات کنندگان افراد را تحت تأثیر قرار بدهد. به‌نحوی که رفتار و تصمیم‌گیری افراد تحت تأثیر مستقیم ایشان باشد (مولر، ۲۰۲۱الف). این مقوله، هم چالشی جدی در نظام اخلاقی و هم دارای بار حقوقی است. به‌خصوص از جهت مسئولیت مدنی و تسبیب در آسیب‌رسانی و ایجاد صدمات و خسارات گوناگون. مواد ۳۲۸ و ۳۳۱ تا ۳۳۵ قانون مدنی و مواد ۱ و ۸ مسئولیت مدنی و مواد ۵۲۴، ۵۲۵ و به‌خصوص ۵۲۶ قانون مجازات اسلامی مصوب ۱۳۹۲ در مورد سببیت و تسبیب شاهد بر این معناست.

۲-۱-۵. سیستم‌های خودمختار در رویکردی فلسفی - اخلاقی، و فقهی - حقوقی

سیستم‌های خودمختار می‌توانند در مقابل رویدادهای پیش‌بینی نشده بدون کنترل و دخالت مستقیم انسانی تغییر رفتار داشته باشند. این سیستم‌ها حتی در محیط‌های پیچیده می‌توانند با تلفیق الگوریتم‌های هوش مصنوعی، پردازش داده‌ها و حسگرها

¹ Data ethics

² Big Data ethics

³ Machine Learning

⁴ Shallow

⁵ Standard

⁶ Deep

عمل کنند. دیوید پی واتسون^۱ و دیوید اچ شایدت^۲ معتقدند یکی از ویژگی‌های سیستم‌های خودمختار انطباق‌پذیری است به این معنی که می‌توانند در محیط‌های پیش‌بینی‌نشده تعدیل رفتار داشته باشند. همچنین بر اساس نظر آن‌ها خودمختاری دارای دو درجه نیمه مختار (نیازمند دخالت کم انسان)، و کاملاً خودمختار است. سیستم‌های خودمختار خود با چالش‌های متعددی همچون عدم اطمینان از عملکرد سیستم در شرایط غیرمنتظره، عدم اطمینان از هماهنگی بین چند سیستم خودمختار با چالش‌های اخلاقی و سیاسی روبرو است (واتسون و شایدت، ۲۰۰۵)^۳ و بی‌تردید از جهت مسئولیت اجتماعی و فرایند اثرگذاری بر زندگی افراد جامعه و دست‌کم در زمینه‌سازی برای آسیب‌های فراوان اجتماعی و پیش‌گیری از آن، با ابعاد گوناگون حقوقی روبه‌رو و از چالش‌های مهم در حوزه حقوق فن‌آوری‌های نوین اجتماعی، به‌خصوص با نگاه سیستم‌های اجتماعی- فنی است. در نظام فقهی- حقوقی وثوق و اطمینان از شرایط اقدامات مسئولانه است. اقدامات شخص مسئول زمانی واجد وجهت قانونی و فقهی- حقوقی است که بر اساس وثوق و اطمینان رقم بخورد. پزشک در اقدام به معالجه زمانی از عهده مسئولیتش برمی‌آید و در صورت ایجاد هرگونه خسارت و آسیبی، نه تنها از اتهام عمد در جنایت نجات می‌یابد، بلکه حتی از ضمانت دیه هم رهایی می‌یابد که اقداماتش بر اساس وثوق و علم و آگاهی باشد و قصور و تقصیری در علم و عمل نداشته باشد. ماده ۴۹۵ قانون مجازات اسلامی مصوب ۱۳۹۲، به‌خصوص در تبصره ۱ بر این مطلب به‌صراحت تأکید دارد.

بنابراین، سیستم‌های خودمختار از این جهت با چالش فقه- حقوقی جدی روبه‌روست و در صورت تحصیل وثوق و اطمینان در عملکرد سیستم‌های هوش مصنوعی خودمختار می‌توان آن عملکرد را به رسمیت حقوقی شناخت. عملکرد اجتهاد در استنباط احکام و یا عملکرد قضائی سیستم‌های هوشمند مانند COMPAS در آمریکا (ویشمایر، ۲۰۲۰، ۳۰۲) و HART در بریتانیا (اسکانتامبورلو، ۲۰۱۹، ۵۷-۵۸) همگی بر این اساس استوار است.

از نظر فلسفی نیز با توجه به جایگاه و ارزش علم و وثوق در کلیه مراحل وجودی (ملاصدرا، ۱۳۶۸، ۲، ۲۳۵) نمی‌توان عملکرد سیستم‌های هوشمند فاقد وثاقت و اطمینان در رفتار و عملکرد را واجد وجهت فلسفی دانست.

ارزش وثاقت و اطمینان در عملکرد از خاستگاه هر دو وجه عرف عام و انواع عرف‌های خاص برخوردار است و از این رهگذر می‌توان بر این امر تأکید داشت که این مقوله چه از نظر فلسفی و چه از منظر فقهی- حقوقی در مدل اندیشگی سیستم‌های اجتماعی- فنی نیز مورد تأیید و تأکید است.

۲-۱-۶. اتوماسیون و اشتغال در رویکردی فلسفی- اخلاقی، و فقهی- حقوقی

شاید این مورد بیشترین دغدغه در روی کار آمدن سیستم‌های هوش مصنوعی برای بشر باشد و تصویری که بسیاری از افراد دارند این است که بسیاری از مشاغل با روی کار آمدن هوش مصنوعی از دست خواهد رفت (مولر، ۲۰۲۱الف).

چنین واقعیت اجتماعی زمینه این پرسش حقوقی را فراهم می‌آورد که در حکمرانی خوب، مسئولیت تأمین مشاغل افراد در حوزه‌های مورد تعرض از سوی ماشین‌های هوشمند و حمایت‌های حقوقی و اقتصادی- اجتماعی از افراد آسیب‌پذیر در این فرایند آیا بر عهده حاکمیت و از وظائف دولت به شمار می‌رود؟! و اساساً در چنین شرایطی نقش حکمرانی خوب از چه الزامات حقوقی متأثر است؟!

^۱ David P Watson

^۲ David H Scheidt

^۳ در مورد ربات‌های خود مختار نیز، در کتاب مقدمه‌ای به ربات‌های خود مختار؛ پیش درآمد جامعی به مفاهیم پایه‌ای رباتیک و رباتیک خود مختار ارائه شده است. بیشترین توجه به الگوریتم‌ها و روش‌هایی است که به ربات‌ها اجازه می‌دهد بدون کنترل و دخالت انسان عمل کنند در این کتاب به نحو آموزشی و قابل فهم برای همگان مبانی رباتیک خود مختار، ادراک محیط، برنامه ریزی حرکت، یادگیری ماشین، چالش‌ها و محدودیت‌ها و آورده شده است (واهد، ۲۰۱۶).

از نگاه فلسفه اخلاق نیز این امر مهم است که دانسته شود دال مرکزی و عامل اخلاقی در ارتباط با این وضعیت اجتماعی - حقوقی چگونه تعریف و ترسیم گردد.

۲-۱-۷. تکینگی در رویکردی فلسفی - اخلاقی، و فقهی - حقوقی

سطح هوش مصنوعی ممکن است به ابر بشری برسد. مولر و کنون استدلال‌های مرتبط با خطر وجودی^۱ هوش مصنوعی فوق هوشمند را بررسی کرده‌اند و معتقدند بر دو مقدمه قضیه متعامد^۲ و قضیه تکینگی^۳ قرار دارد^۴. در تکینگی اذعان می‌شود که هوش مصنوعی فوق هوشمند از کنترل انسان خارج می‌شود و این یک خطر وجودی است (عاشوری و پرویزی، ۱۴۰۳، ۱۳۸). چالش‌های فلسفی - اخلاقی این مسأله به خوبی قابل لمس است. آموزه‌های اخلاقی و فلسفه اخلاق در بستری شناور در فرهنگ بشری رشد و توسعه یافته و واجد تعامل با حوزه‌های مختلف است، ولی در سیستم‌های هوش مصنوعی در سطح ابر بشری به کدام تضمین و اطمینان می‌توان از این جهت خیالی آسوده داشت؟! همچنین در اینجا یکی از چالش‌های جدی حقوقی، بلکه نظام جدید حقوقی با عناوینی مثل حقوق ربات‌های هوشمند شکل می‌گیرد و آثار علمی گوناگونی در این خصوص رقم می‌خورد.^۵ و باید هر چه سریع‌تر در جهت تدوین نظام واحد بین‌المللی در این مقام و برای این مرحله کوشید تا بر عنصر شمولگرایی در سطح جهانی صیانت شود و مقتضای ذاتی فناوری هوش مصنوعی را که در گریز از هر قید و بند و محدودیتی است و بر پایه کلان داده‌ها می‌تواند به رشد و توسعه مضاعف دست یابد تأمین کرد.

۲-۲. اخلاق ماشین و فلسفه اخلاق

شاید سخن از اخلاق ما را به زمان سقراط برگرداند. «ونگ» اخلاق را مجموعه واحدی می‌داند که برای دفع تعارضات درونی و بیرونی به میان آمده‌اند. اخلاق را می‌توان به‌طور کلی پاسخ به سؤالات در مورد فضیلت، عدالت، زندگی درست و ... دانست (ایمان و غفاری‌نسب، ۱۳۹۰، ۴). در این جا مجال بحث از شاخه‌های مختلف فلسفه اخلاق نیست. ما باید به اخلاق ماشین به‌عنوان یکی از چالش‌های پیش روی هوش مصنوعی بپردازیم. اما حائز اهمیت است که بدانیم چنین تعاریفی از اخلاق و ساخت فلسفه‌ای از این سنخ برای نظام اخلاقی نمی‌تواند همان چیزی باشد که در یک سیستم اجتماعی - فنی مورد توجه است. چون همانگونه که در تعریف سیستم‌های اجتماعی - فنی گذشت مدار همه بایسته‌ها؛ چه بایسته‌های اخلاقی و چه بایسته‌های فقهی - حقوقی مبتنی بر اقتضائات عرفی و مقتضای اجتماعی در تعامل با اقتضائات فناوری هوش مصنوعی شکل می‌گیرد و نمی‌توان با تعاریف و الزامات از پیش ساخته‌ای ضرورت‌های سیستم‌های اجتماعی - فنی را در فناوری‌های نوین تأمین و تضمین کرد.

مسئله مورد توجه، تضمین رفتار اخلاقی ماشین‌هاست و اینکه آیا رفتار آن‌ها از نظر انسان‌ها مورد قبول است یا خیر؟ در مورد اخلاق ماشین‌ها عده‌ای معتقدند که ماشین اساساً نمی‌تواند اخلاق داشته باشد زیرا فاقد احساس است و برخلاف آن‌ها عده‌ای دیگر بر این باور هستند که انسان‌ها خود یک ماشین بوده و دارای حالت‌های اخلاقی‌اند و به همین منوال ماشین‌ها دارای حالت‌های اخلاقی‌اند (رمضانی و فیضی درخشی، ۱۳۹۲، ۳).

در سال ۱۹۴۲ آیزاک آسیموف نویسنده داستان‌های علمی - تخیلی سه قانون در جهت آسیب نرساندن ربات‌ها به انسان‌ها مطرح کرد:

^۱ Existential Risk

^۲ Orthogonality

^۳ Singularity

^۴ رک. (مولر و کنون، ۲۰۲۲)

^۵ رک. اوگوپاگالو، ۱۳۹۹، حقوق ربات‌ها، ترجمه پیمان نمایان، نشر میزان

۱. ربات‌ها نباید با انجام دادن یا ندادن عملی به انسان‌ها آسیب برسانند.

۲. ربات‌ها موظف‌اند به فرمان‌های انسان‌ها عمل کنند مگر اینکه به انسان‌ها آسیب برسانند.

۳. ربات باید تا جایی که دو قانون مطرح شده رعایت می‌شود، به حفظ بقای خویش ملزم باشد.

روشن است که سه اصل اخلاقی آسیموف و همچنین هر نظم اخلاقی دیگری در این قلمرو، اخلاقی انتزاعی و فرا حقوقی نیست و به تأکید رابطه مستقیم با چینش نظام حقوقی مطلوب در حوزه هوش مصنوعی و ربات‌های هوشمند دارد، چنانکه تضمین حقوق بشر و حقوق اساسی انسان‌ها و همچنین حقوق محیط‌زیست در این زمینه یکی از جدی‌ترین و ضروری‌ترین اصولی است که باید مورد توجه عمیق واقع شود.

در مورد هوش مصنوعی، سیستم‌ها باید به‌سوی ماشین‌های انتخابی‌ای بروند که طراحی آن‌ها متوجه ارزش‌های انسانی باشد و با توجه به اینکه بعضی از دلایل ادراکات درونی و اندیشه‌های اخلاقی انسان‌ها هنوز ناشناخته‌اند؛ برای طراحی این سیستم توجه به محدودیت‌های ابزاری واجب است (رضانی و فیضی درخشی، ۱۳۹۲، ۷). از سوی دیگر با توجه به گسترش بهره‌گیری از هوش مصنوعی تقریباً در همه ابعاد جهانی و حمایت گسترده از این تکنولوژی، جامعه جهانی را به‌سوی تدوین قواعدی در تحقیق، تولید، عملکرد و ... هوش مصنوعی سوق داد تا پاسخگوی چالش‌های مطرح و احتمالی در این زمینه باشد. یکی از این تلاش‌ها، تدوین پیش‌نویس یونسکو در اخلاق هوش مصنوعی در سپتامبر ۲۰۲۰ است. در این توصیه‌نامه به چهار ارزش بنیادین انسانی شامل احترام، ارتقا و تضمین حقوق اساسی بشر، حفاظت از محیط‌زیست و اهمیت زندگی در صلح اشاره شده است و از تمامی فعالان در تمامی عرصه‌های هوش مصنوعی خواسته شده به این اصول و ارزش‌ها پایبند باشند (سیفی و رزمخواه، ۱۴۰۲، ۱۴ و سایت یونسکو)^۱. اما با وجود ابهامات و عدم صراحت بعضی از مفاد این توصیه‌نامه هنوز نتوانسته است به‌عنوان قانون کلی و اساسنامه در اخلاق هوش مصنوعی مورد قبول همگان باشد و همه جوامع بشری و افراد و زیست‌بوم‌ها را پوشش دهد.

۲-۳. لزوم شمول‌گرایی اخلاقی - حقوقی در هوش مصنوعی

درست است که قانون ابزار قدرتمندی است اما توجه به اخلاق است که در ایجاد هوش مصنوعی ایمن و سودمند ضروری می‌نماید. قانون جامع است اما بسیاری از حیطه‌های زندگی بشر را در بر نمی‌گیرد. همچنین قوانینی وجود دارند که علیرغم منطبق بودن با معیارهای ساکنان کشورهای وضع‌کننده، هنوز غیراخلاقی هستند. پس نمی‌توان قانون را مشروع‌ترین نظام قابل تصور دانست (هندریکز، ۲۰۲۵، ۲۸۹).

تلاش برای دستیابی به یک نظام اخلاقی که مورد قبول همگان و بدون پیش‌فرض و سوگیری باشد تاکنون بی‌نتیجه بوده است و ما در طول زمان با اقسام نظام‌های اخلاقی بر پایه‌های اهداف مختلف روبرو بوده‌ایم که فلسفه‌های اخلاقی گوناگونی آن‌ها را پی‌ریزی کرده است. مکس تگمارک معتقد است علی‌رغم عدم اجماع اخلاقی بین جوامع بشری می‌توان اصولی را به‌عنوان پایه و اساس یافت که درباره آن اتفاق نظر وجود دارد. به نظر می‌رسد باید تلاش کرد ظرفیتی مورد وفاق همگان را در نظامی اخلاقی - حقوقی برای صیانت از جامعه بشری در سطح بین‌المللی در همه ابعاد حقوقی و اخلاقی فراهم آورد. این ضرورت و الزام البته در سیستم‌های اجتماعی فنی بر پایه نظام یکپارچه عوامل گوناگون اجتماعی در کنار اقتضائات فنی فن‌آوری‌های هوش مصنوعی باید شکل بگیرد و نمی‌تواند به‌صورت جهان‌های جداگانه به دنیای حقوق و اصول حقوقی و به جهان اخلاق و بایسته‌های اخلاقی نظر داشت. تگمارک ۴ اصل را اساسی می‌داند:

۱. فایده‌باوری به این معنی که تجربیات مثبت بیشتر و رنج کمتر شود؛

^۱ <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>

۲. تنوع، اینکه تجربه‌های مثبت متنوع از تجربه‌ای واحد که مکرراً اتفاق می‌افتد حتی اگر تجربه‌ای مثبت باشد، بهتر است؛

۳. خود آیینی^۱ هم موجودات و هم جوامع در پیروی اهدافشان تا جایی که با اصل بنیادی‌تری در تضاد نباشد باید آزاد باشند؛

۴. میراث، سازگاری با سناریوهایی که اکثریت انسان‌های امروزی خوشایند تلقی می‌کنند و عدم سازگاری با آنچه تقریباً تمام انسان‌های امروزی وحشتناک می‌دانند.

وی در مورد فایده باوری برخلاف عموم که آن را «بالاترین سعادت برای بیشترین تعداد مردم می‌دانند» (انسان‌محوری)، آن را به همه موجودات حتی غیرانسانی و ذهن‌های بشری شبیه‌سازی‌شده آگاه و دیگر موارد هوش مصنوعی در آینده تعمیم می‌دهد. وی تعریف خود را طبق آنچه گفته شد بر تربیت استوار می‌کند چون معتقد است بدون تجربه، معنای مرتبط با اخلاق وجود ندارد. با تکیه بر همین اصل می‌توان لزوم فایده باوری برای همه ذی‌نفعان در سیستم‌های هوش مصنوعی را مبنا قرار داد. اما ضرورت تأمین نظام حقوقی جامع در جهت صیانتی بین‌المللی از حقوق بشر و حقوق اساسی انسان‌ها و حقوق محیط‌زیست نباید با کمترین غفلتی روبه‌رو گردد.

۳. فقه و حقوق و شمول‌گرایی در هوش مصنوعی

قانون بدون تفسیر انسانی ابزاری ناکامل است. همواره شکافی میان قواعد مكتوب و واقعیت‌های پیچیده عینی زندگی وجود دارد که نیازمند قضاوت و تفسیری است که خود می‌تواند محل مناقشه باشد. بنابراین حتی اگر هوش مصنوعی منطبق با قانون تنظیم شود، نقش قضاوت انسانی در فرایند حقوقی اجتناب‌ناپذیر است. زیرا مشروعیت و اعتبار قانون تنها زمانی است که تفسیر آن بر آمده از نمایندگان مردم باشد یا دست کم تحت نظارت و ارزیابی‌های دموکراتیک قرار بگیرد (هندریکز، ۲۰۲۵، ۲۸۹).

باید توجه داشت که اینکه قانون مشروعیت داشته باشد، نشانگر اخلاقی بودن آن نیست. حتی اگر قوانین به‌صورت دموکراتیک تدوین شوند ممکن است بعدها به‌عنوان خطایی تاریخی از نظر اخلاقی (برای مثال ممنوعیت مهاجرت که امروز چالشی برای بسیاری جوامع است یا جنگ‌های نامشروع، برده‌داری قانونی و نژادپرستی نهادینه‌شده در جامعه) شناخته شوند. همچنین باید در نظر داشت همواره قوانین در نسبت با تحولات اخلاقی جامعه عقب‌تر هستند و گاه بسیار به طول می‌انجامد که قوانین از ارزش‌های اخلاقی جوامع حمایت و پیروی کنند (جان جی، ۲۰۲۳)، ولی این نقیصه نمی‌تواند توجیهی برای فراموشی ضرورت تدوین نظام جامع حقوقی در سطح بین‌المللی در حوزه هوش مصنوعی و ربات‌های هوشمند قلمداد شود و الا بیشترین آسیب به حقوق اساسی انسان‌ها و حتی به محیط‌زیست وارد خواهد آمد.

نتیجه‌گیری

برآیند پژوهش حاضر در دو بعد به‌شدت ضروری اخلاق و حقوق در سطح جامعه بین‌المللی در راستای توجه به شمول‌گرایی در هوش مصنوعی بر این معنا استوار است که باید برای رعایت جنبه عام و شمول‌گرایی بخصوص بر اساس سیستم‌های اجتماعی-فنی به آگاهی دست یافت که زمینه‌های سوگیری و تبعیض و نقض حریم خصوصی و نادیده گرفتن اخلاق تجربه‌پذیر و کارآمد را در جهان غرق در هوش مصنوعی و سرشار از ربات‌های هوشمند فراهم آورد. برای دستیابی به نظام اخلاقی جامع و واجد شمول‌گرایی باید بر پایه فلسفه‌های گوناگون نظام اخلاقی قابل تعمیم و پذیرشی را شکل داد. در جنبه حقوقی نیز توجه به مبانی

¹ autonomy

و قواعد فقه امامیه در حوزه حقوق داخلی می‌تواند خصوصیت شمولگرایی را با بن‌مایه‌های مستحکم‌تر حقوقی و عرفی همراه کند و اینگونه از منظر فلسفی- اخلاقی و فقهی- حقوقی ویژگی شمول‌گرایی را در برنامه‌ای جامع در سیستم‌های هوش مصنوعی جستجو کرد.

برای دستیابی به این منظور به‌غایت واجد اهمیت یکی از گام‌های ضروری آن است که از حصار مفهوم شهروندی باید فراتر رفت و به جایگزینی عاملی جهان‌شمول باید عطف توجه کرد. چون غیر شهروندان ذی‌نفع نیز در سیطره ماشین‌های هوشمند و جهان هوش مصنوعی قرار دارند و تکیه بر صرف عنصر شهروندی در چنین جهان پیچیده و حساسی نه‌تنها تضمینی در جهت حقوق بشر و حقوق اساسی انسان‌ها نیست، بلکه خود زمینه‌ای برای انواع تبعیض‌ها و سوگیری‌ها و نقض حریم خصوصی افراد و ده‌ها مورد غیراخلاقی و ضد اصول حقوقی دیگر است تا چه رسد به هویت سیستم‌های اجتماعی- فنی که صرف‌نظر از این تبعات برای نفس شمول‌گرایی ارزش عرفی- اجتماعی به اقتضای ذات فناوری قائل است و باید از این جهت نیز شمول‌گرایی را مطمح نظر قرار داد. لذا شاید عنوان «فرد» بجای شهروند بتواند در این راستا گزینه‌ای پیشنهادی باشد. اما به نظر گزینه‌ای کافی و کارآمد نیست. چون شمولش نسبت به شخصیت‌های حقوقی کمی سخت به نظر می‌رسد. بنابراین، گزینه «شخص و اشخاص» به جهت شمولش بر اشخاص حقیقی و حقوقی در سطح عموم، اعم از شهروندان و غیر آن‌ها می‌تواند پیشنهاد مناسبی باشد. اما باید توجه داشت در سیستم اجتماعی- فنی چون به مجموعه عوامل مختلف اجتماعی در کنار اقتضات گوناگون فنی یکجا به‌عنوان زنجیره‌ای جامع از عوامل گوناگون یکپارچه در جهان فن‌آوری توجه می‌شود و چنین نیست که زمینه‌های اخلاقی به‌طور مجزی و الزامات حقوقی به‌طور جداگانه ملاحظه گردد، پس باید اخلاق و حقوق را در همان عوامل گوناگون اجتماعی و زمینه‌های متعدد اجتماعی در کنار اقتضات مختلف و متفاوت جهان فن و فن‌آوری به‌عنوان یک زنجیره واحد به‌هم‌پیوسته دید.

منابع

- ابن سینا. (۱۳۸۶). شرح‌الاشارات، چاپ ۱، قم: بوستان کتاب.
- اسدپور، فرشته؛ پهلوان زاده، عباس و خدانی، پدram. (۱۴۰۲). «تأثیر تحولات هوش مصنوعی در حوزه حقوق مالکیت فکری»، تحقیقات حقوق خصوصی و کیفری، (۵۸)، ۱۳۱-۱۴۷. <https://www.noormags.ir/view/fa/articlepage/2178994>
- امینی، اعظم. (۱۳۹۵). «بازاندیشی رابطه انصاف و حقوق»، فصلنامه پژوهش حقوق خصوصی، ۵ (۱۶)، ۵۶-۳۱. <https://doi.org/10.22054/jplr.2016.6740>
- بجنوردی، سیدحسن. (۱۳۷۷). القواعد الفقهیه، چاپ ۱، قم: نشر الهادی.
- بروجردی، حسین. (۱۴۱۳ق). تقریرات ثلاثه، مقرر: علی پناه اشتهااردی، چاپ ۱، قم: جامعه مدرسین.
- ثقه الاسلام الکلینی، محمد بن یعقوب. (۱۳۶۳). الکافی، چاپ ۵، تهران: دارالکتب الاسلامیه.
- حسینی، محمد جواد. (بی تا). مفتاح الکرامه، چاپ ۱، بیروت: دار احیاء التراث العربی.
- خوئی، سید ابوالقاسم. (۱۴۱۸ق). موسوعه الامام الخوئی، چاپ ۱، قم: مؤسسه احیاء آثار الامام الخوئی.
- رمضانی، مجید. فیضی درخشانی، محمدرضا (۱۳۹۲). اخلاق ماشین: چالش‌ها و رویکردهای مسائل اخلاقی در هوش مصنوعی و ابرهوش. فصلنامه اخلاق در علوم و فناوری، ۸ (۴)، ۳۵-۴۳. <https://dorl.net/dor/20.1001.1.22517634.1392.8.4.4.7>
- سبحانی، جعفر. (۱۳۹۷). الموجز، چاپ ۲۷، قم: مؤسسه امام صادق علیه السلام.
- سیفی، آناهیتا. رزمخواه، نجمه. (۱۴۰۲). «نقدی بر پیش‌نویس توصیه‌های یونسکو در اخلاق هوش مصنوعی از منظر حق بر محیط زیست سالم». پژوهش حقوق عمومی، ۲۴ (۷۸)، ۹-۴۷. <https://doi.org/10.22054/qjpl.2022.63030.2659>
- شیخ انصاری، مرتضی. (۱۴۱۵ق). کتاب‌المکاسب، چاپ ۱، قم: کنگره شیخ.
- شیخوند، محمدصادق؛ کردعلیوند، روح‌الدین؛ مینائی، بهروز؛ آشوری، محمد و مهدوی ثابت، محمد علی. (۱۴۰۲). «هوش مصنوعی و صدور احکام کیفری؛ تصمیم‌سازی یا تصمیم‌گیری؟»، پژوهش‌های حقوق تطبیقی، ۲۷، (۴)، ۱۶۷-۱۳۸.
- صفدری شریانی، ابوطالب. (۱۴۰۳). هوش مصنوعی: فلسفه، اخلاق و جامعه، تهران: پژوهشگاه فضای مجازی.

- طباطبایی، محمدحسین. (۱۳۶۸). *اصول فلسفه و روش رئالیسم*، چاپ ۶، قم: انتشارات صدرا.
- طباطبایی، محمدحسین. (بی‌تا). *بدایه الحکمه*، قم: جامعه مدرسین.
- عاشوری کیسمی، محمدعلی و مریم پرویزی. (۱۴۰۳). «بررسی عاملیت اخلاقی هوش مصنوعی عام». *پژوهش‌های علم و دین*، ۱۵(۱): ۱۲۵-۱۵۱. <https://doi.org/10.30465/srs.2024.49282.2162>
- عاشوری کیسمی، محمدعلی. (۱۴۰۳ الف). «طرح و بررسی برخی از مسائل اخلاقی هوش مصنوعی در هنر». *متافیزیک*، ۱۶(۳۷)، ۹۳-۱۱۰. <https://doi.org/10.22108/mp.2024.138105.1488>
- عاشوری کیسمی، محمدعلی. (۱۴۰۳ ب). «همگرایی حریم خصوصی و شفافیت، محدودیت‌های طراحی هوش مصنوعی». *حکمت و فلسفه*، ۲۰(۷۸)، ۷۳-۴۵. <https://doi.org/10.22054/wph.2024.75680.2183>
- عاشوری کیسمی، محمدعلی و مریم پرویزی و سیدمحمدحسن مومنی. (۱۴۰۴). «پراگماتیسم، اخلاق هوش مصنوعی و سیستم‌های اجتماعی-فنی» *مجله پژوهش‌های فلسفی*، ۱۹(۵۲)، ۲۸۱-۳۰۰. <https://doi.org/10.22034/jpiut.2025.66643.4060>
- عباسی، محمود؛ تیموری، مهرداد. (۱۴۰۲). «مروری بر چالش‌های اخلاقی و حقوقی کاربرد هوش مصنوعی در نظام سلامت». *مجله اخلاق پزشکی*، ۱۷(۴۸)، ۱-۱۱. <https://doi.org/10.22037/mej.v17i48.44053>
- قرطبی، محمد بن احمد. (۱۳۸۴ ق). *الجامع لأحكام القرآن*، چاپ ۲، قاهره: دار الکتب المصریه.
- گروه متخصصان یونسکو. (۱۴۰۱). *هوش مصنوعی و نظام آموزشی*، مترجم: کاوه بازرگان، تهران: کند و کاو.
- محقق اصفهانی، محمد حسین. (۱۴۲۹ ق). *نهایه الدرایه*، چاپ ۲، بیروت: مؤسسه آل البيت عليهم السلام لإحياء التراث.
- مظفر، محمد رضا. (۱۳۸۷). *اصول الفقه*، چاپ ۵، قم: بوستان کتاب.
- مکارم شیرازی، ناصر. (۱۳۸۵). *دائرة المعارف فقه مقارن*، چاپ ۱، قم: مدرسه امام علی بن ابی طالب علیه السلام.
- ملاصدرا. (۱۳۶۸). *الحکمه المتعالیه*، چاپ ۱، قم: مکتبه المصطفوی.
- میرداماد، محمدباقر. (۱۳۹۳). *ضوابط الرضا*، چاپ ۲، قم: مرکز نشر المصطفی.
- نجفی، محمدحسن. (۱۴۲۱ ق). *جواهر الکلام*، چاپ ۱، قم: موسسه دائرة المعارف فقه مقارن.
- هاشمی، سید محمود و دیگران. (۱۳۸۲). *فرهنگ فقه مطابق مذهب اهل بیت عليهم السلام*، چاپ ۱، قم: مؤسسه دائرة المعارف فقه اسلامی.

References

- Abbasi, M., & Teymouri, M. (2023). A review of the ethical and legal challenges of applying artificial intelligence in the health system. *Journal of Medical Ethics*, 17(48), 1–11. <https://doi.org/10.22037/mej.v17i48.44053> (in persian)
- Amini, A. (2016). Rethinking the relationship between fairness and law. *Private Law Research Quarterly*, 5(16), 31–56. <https://doi.org/10.22054/jplr.2016.6740> (in persian)
- Angelo, J. A., Jr. (2006). *Robotics: A reference guide to the new technology* (1st ed.). Greenwood Press.
- Asadpour, F., Pahlavanzadeh, A., & Khandani, P. (2023). The impact of developments in artificial intelligence in the legal field of intellectual property. *Private and Criminal Law Research*, (58), 131–147. <https://www.noormags.ir/view/fa/articlepage/2178994/> (in persian)
- Ashuri Qaismi, M. (2024a). A study of some ethical issues of artificial intelligence in art. *Metaphysics*, 16(37), 93–110. <https://doi.org/10.22108/mp.2024.138105.1488> (in persian)
- Ashuri Qaismi, M. (2024b). The convergence of privacy and transparency, limitations of artificial intelligence design. *Wisdom and Philosophy*, 20(78), 45–73. <https://doi.org/10.22054/wph.2024.75680.2183> (in persian)
- Ashuri Qaismi, M., & Parvizi, M. (2024). A study of the ethical agency of general artificial intelligence. *Science and Religion Studies*, 15(1), 125–151. <https://doi.org/10.30465/srs.2024.49282.2162> (in persian)

- Ashuri Qismi, M., Parvizi, M., & Momeni, S. M. (2025). Pragmatism, ethics of artificial intelligence and socio-technical systems. *Journal of Philosophical Research*, 19(52), 281–300. <https://doi.org/10.22034/jpiut.2025.66643.4060> (in persian)
- Bakiner, O. (2014). *Paths, dangers, strategies*. Oxford University Press.
- Bakiner, O. (2023). Pluralistic sociotechnical imaginaries in Artificial Intelligence (AI) law: The case of the European Union's AI Act. *Law, Innovation and Technology*, 15(2), 558–582. <https://doi.org/10.1080/17579961.2023.2245675>
- Bojnardi, S. H. (1998). *Al-qawwaad al-fiqhiyyah* (1st ed.). Nashr al-Hadi, Qom, Iran. (in persian)
- Boroujerdi, H. (1992). *Three editions* A. P. Eshtehardi, Ed. 1st ed, Jamia Madrasain, Qom, Iran. (in persian)
- Bostrom, N. (2012). The superintelligent will: Motivation and instrumental rationality in advanced artificial agents. *Minds & Machines*, 71–85.
- Coeckelbergh, M. (2020). *AI ethics*. The MIT Press.
- Coeckelbergh, M., & Sætra, H. S. (2023). Climate change and the political pathways of AI: The technocracy-democracy dilemma in light of artificial intelligence and human agency. *Technology in Society*, 75, Article 102406. <https://doi.org/10.1016/j.techsoc.2023.102406>
- Dignum, V. (2018). Ethics in artificial intelligence: Introduction to the special issue. *Ethics and Information Technology*, 20(1), 1–3. <https://doi.org/10.1007/s10676-018-9450-z>
- Ebers, M., Hoch, V. R., Rosenkranz, F., Ruschemeier, H., & Steinrötter, B. (2021). The European Commission's proposal for an Artificial Intelligence Act—A critical assessment by members of the Robotics and AI Law Society (RAILS). *J*, 4(4), 589–603. <https://doi.org/10.3390/j4040043>
- Engstrom, D. F., & Haim, A. (2023). Regulating government AI and the challenge of sociotechnical design. *Annual Review of Law and Social Science*, 19(1), 277–298.
- Floridi, L. (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
- Greenstein, S. (2022). Preserving the rule of law in the era of artificial intelligence (AI). *Artificial Intelligence and Law*, 30(3), 291–323. <https://doi.org/10.1007/s10506-021-09294-4>
- Hashemi, S., et al. (2003). *The dictionary of fiqh according to the school of the Ahl al-Bayt (peace be upon them)* (1st ed.). Institute of the Encyclopedia of Islamic Fiqh, Qom, Iran. (in persian)
- Hendrycks, D. (2025). *Introduction to AI safety, ethics, and society*. CRC Press. <https://doi.org/10.1201/9781003530336>
- Holton, R., & Boyd, R. (2021). “Where are the people? What are they doing? Why are they doing it?” (Mindell): Situating artificial intelligence within a socio-technical framework. *Journal of Sociology*, 57(2), 179–195. <https://doi.org/10.1177/1440783319873046>
- Hopster, J. (2021). What are socially disruptive technologies? *Technology in Society*, 67, Article 101750. <https://doi.org/10.1016/j.techsoc.2021.101750>
- Hussaini, M. (n.d.). *Miftah al-karama*, 1st ed, Dar Ihya al-Turat al-Arabyyah, Beirut, Lebanon. (In Arabic)
- Ibn Sina. (2007). *Sharh al-isharat*, 1st ed, Bostan Ketab, Qom, Iran. (in persian)
- Iman, M. T., & Ghafari Nasab, E. (2011). Ethical criteria in humanities research. *Ethics in Science and Technology Quarterly*, (2), 4. (in persian)
- Johnson, C. Y. (2019, October 24). Racial bias in a medical algorithm favors white patients over sicker black patients. *The Washington Post*. <https://www.washingtonpost.com/health/2019/10/24/racialbias-medical-algorithm-favors-white-patients-over-sicker-black-patients/>
- Khoi, S. A. (1997). *Musoo 'ah al-imam al-khoi*, 1st ed, Ihya'a at-Ta'r al-Imam al-Khoi, Qom, Iran. (in persian)

- Kudina, O., & van de Poel, I. (2024). A sociotechnical system perspective on AI. *Minds and Machines*, 34(3), Article 21. <https://doi.org/10.1007/s11023-024-09680-2>
- Makarem Shirazi, N. (2006). *Encyclopedia of contemporary jurisprudence*, 1st ed, Madrasa Imam Ali ibn Abi Talib, Qom, Iran. (in persian)
- Mirdamad, M. (2014). *The rules of breastfeeding*, 2nd ed, Al-Mustafa Publishing Center, Qom, Iran. (in persian)
- Mohaqqeq Isfahani, M. H. (2008). *The end of al-diriyah*, 2nd ed, Aal-Bayt Foundation for the Revival of Heritage, Beirut, Lebanon. (In Arabic)
- Mozaffar, M. R. (2008). *Principles of jurisprudence*, 5th ed, Bostan Ketab, Qom, Iran. (in persian)
- Mulla Sadra. (1989). *Transcendental wisdom*, 1st ed, Maktaba al-Mustafawi, Qom, Iran. (in persian)
- Muller, V. (2021a). Ethics of artificial intelligence and robotics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved December 18, 2024, from <https://plato.stanford.edu/entries/ethics-ai/>
- Muller, V. (2021b). Deep opacity undermines data protection and explainable artificial intelligence. In *Overcoming opacity in machine learning*.
- Muller, V., & Cannon, M. (2022). Existential risk from AI and orthogonality: Can we have it both ways? *Ratio*, 25–36. <https://doi.org/10.1111/rati.12320>
- Najafi, M. H. (2000). *Jawaher al-kalam*, 1st ed, Institute of the Encyclopedia of Contemporary Fiqh, Qom, Iran. (in persian)
- Nay, J. J. (2023). Law informs code: A legal informatics approach to aligning artificial intelligence with humans. *Northwestern Journal of Technology and Intellectual Property*, 20(3). arXiv:2209.13020
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28(4), 4221–4241. <https://doi.org/10.1007/s10639-022-11316-w>
- Nyholm, S., Friedman, C., Dale, M. T., Puzio, A., Babushkina, D., Löhr, G., Gwagwa, A., Kamphorst, B. A., Perugia, G., & IJsselsteijn, W. (2023). Social robots and society. In I. Van de Poel, L. E. Frank, J. Hermann, J. Hopster, D. Lenzi, S. Nyholm, B. Taebe, & E. Ziliotti (Eds.), *Ethics of socially disruptive technologies* (pp. 53–82). Open Book. <https://doi.org/10.11647/obp.0366.03>
- Ponkin, I. V., & Redkina, A. I. (2018). Artificial intelligence from the point of view of law. *RUDN Journal of Law*, 22(1), 91–109. <https://doi.org/10.22363/2313-2337-2018-22-1-91-109>
- Qurtubi, M. (1964). *Al-jami' al-ahkam al-qur'an*, 2nd ed, Dar al-Kutb al-Masriya, Cairo, Egypt. (In Arabic)
- Ramezani, M., & Fayzi Derakhshani, M. (2013). Machine ethics: Challenges and approaches to ethical issues in artificial intelligence and superintelligence. *Ethics in Science and Technology Quarterly*, 8(4), 35–43. <https://dorl.net/dor/20.1001.1.22517634.1392.8.4.4.7> (in persian)
- Sadek, M., Calvo, R. A., & Mougenot, C. (2023). Designing value-sensitive AI: A critical review and recommendations for socio-technical design processes. *AI and Ethics*, 4, 949–967. <https://doi.org/10.1007/s43681-023-00373-7>
- Safdari Sharabiani, A. (2024). *Artificial intelligence: Philosophy, ethics and society*. Virtual Space Research Institute, Tehran, Iran. (in persian)
- Saifi, A., & Razmkhah, N. (2023). A critique of UNESCO's draft recommendations on the ethics of artificial intelligence from the perspective of the right to a healthy environment. *Public Law Research*, 24(78), 9–47. <https://doi.org/10.22054/qjpl.2022.63030.2659> (in persian)
- Scantamburlo, T., Charlesworth, A., & Cristianini, N. (2019). Machine decisions and human consequences. In K. Yeung & M. Lodge (Eds.), *Algorithmic regulation*. Oxford University Press.

- Sheikh Ansari, M. (1994). *Kitab al-mukasib*, 1st ed, Sheikh Congress, Qom, Iran. (in persian)
- Sheikhvand, M. S., Kordalivand, R., Minaei, B., Ashouri, M., & Mahdavi Sabet, M. A. (2023). Artificial intelligence and criminal sentencing: Decision making or decision-making? *Comparative Law Research*, 27(4), 138–167. (in persian)
- Shidiq, M. (2023). The use of artificial intelligence-based Chat-GPT and its challenges for the world of education: From the viewpoint of the development of creative writing skills. *Proceeding of International Conference on Education Society and Humanity*, 1(1), 353–357. <https://ejournal.unuja.ac.id/index.php/icesh/article/view/5614>
- Sobhani, J. (2018). *Al-mujiz*, 27th ed, Imam Sadiq (AS) Institute, Qom, Iran. (in persian)
- Stokel-Walker, C. (2023). ChatGPT replicates gender bias in recommendation letters. *Scientific American*. <https://www.scientificamerican.com/article/chatgpt-replicates-gender-bias-in-recommendation-letters>
- Tabatabaei, M. H. (1989). *Principles of philosophy and the method of realism*, 6th ed, Sadra Publications, Qom, Iran. (in persian)
- Tabatabaei, M. H. (n.d.). *The beginning of wisdom* (n.ed.). Jamia Modarresin, Qom, Iran. (in persian)
- Thiqat al-Islam al-Kulaini, M. (1984). *Al-kafi*, 5th ed, Dar al-Kutub al-Islamiyyah, Tehran, Iran. (in persian)
- UNESCO Group of Experts. (2022). *Artificial intelligence and the educational system* (K. Bazargan, Trans.). Kand o Kav, Tehran, Iran. (in persian)
- Van de Poel, I., Frank, L. E., Hermann, J., Hopster, J., Lenzi, D., Nyholm, S., Taebi, B., & Ziliotti, E. (Eds.). (2023). *Ethics of socially disruptive technologies: An introduction*. Open Book. <https://doi.org/10.11647/OBP.0366>
- Wahde, M. (2016). *Introduction to autonomous robots*. Chalmers University of Technology.
- Watson, D. P., & Scheidt, D. H. (2005). Autonomous systems. *Johns Hopkins APL Technical Digest*, 26(4), 368–376.
- Wischmeyer, T., & Rademacher, T. (2020). *Regulating artificial intelligence*. Springer.
- Zhang, J., & Zhang, Z. M. (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC Medical Informatics and Decision Making*, 23(1), Article 7. <https://doi.org/10.1186/s12911-023-02103-9>
- Zhou, K. Q., & Nabus, H. (2023). The ethical implications of DALL-E: Opportunities and challenges. *Mesopotamian Journal of Computer Science*, 2023, 16–21. <https://doi.org/10.58496/MJCSC/2023/003>
- Zowghi, D., & Bano, M. (2024). AI for all: Diversity and inclusion in AI. *AI and Ethics*, 1–4. <https://doi.org/10.1007/s43681-024-00485-8>