

Knowledge without a Subject: a philosophical reflection on the epistemic legitimacy of the machine

Seyyed Sadeqh Zamani 

Assistant Professor, Department of Philosophy, Sa.C., Islamic Azad University, Sanandaj, Iran. E-mail: ss.zamani@iau.ac.ir

Article Info

Article type:

Research Article

Article history:

Received 16 December 2025

Received in revised form 31
December 2025

Accepted 20 January 2026

Published online 18 July 2026

Keywords:

Subjectless Knowledge,
Epistemic Legitimacy, AI
Ethics, Algorithmic
Transparency, Distributed
Responsibility, Philosophy of
Technology

ABSTRACT

The transformation of cognitive and ethical structures in the age of artificial intelligence confronts philosophy with a fundamental question: how does knowledge emerge, and where does responsibility reside when decision-making moves beyond the sphere of human consciousness into algorithmic networks? Focusing on the concept of subjectless knowledge, this study argues that intelligent systems have shifted knowledge from a mental capacity to a mediating process in which human agents, data, and algorithms jointly participate in the production of meaning. Employing a reflective-analytical methodology and examining cases in medicine, media, and law, the paper contends that epistemic validity in the digital era is no longer measured solely by truth, but by the transparency of processes, the capacity for explanation, and the possibility of accountability. Within this framework, social epistemology elucidates how belief is shaped in algorithmic environments, while the ethics of responsibility provides a structure through which the contribution of both human and machinic agents to outcomes can be traced. The proposed model integrates these two dimensions, conceiving knowledge as a mediating event where meaning arises through the interaction between human interpretation and computational reasoning. Accordingly, moral responsibility becomes a distributed property of a network in which every agent participates in the unfolding of cognition. This analysis suggests that maintaining epistemic and ethical legitimacy in intelligent systems requires a philosophical reorientation-from the individual subject toward the distributed architectures of knowing.

Cite this article: Zamani, S. (2026). Knowledge without a Subject: a philosophical reflection on the epistemic legitimacy of the machine. *Journal of Philosophical Investigations*, 20 (55), 541-560. <https://doi.org/10.22034/jpiut.2025.70752.4379>



© The Author(s).

Publisher: University of Tabriz.

Extended Abstract

Introduction and Philosophical Background

Classical Western philosophy has traditionally conceived knowledge as an achievement grounded in the conscious subject. From Cartesian certainty to Kantian transcendental unity, epistemic legitimacy was secured through the reflective capacities of a rational agent capable of intention, judgment, and justification. Knowledge, in this paradigm, presupposed a knowing subject whose awareness constituted the ultimate source of meaning and truth. However, the rise of artificial intelligence and algorithmic decision-making systems has fundamentally unsettled this epistemological architecture. Contemporary societies increasingly rely on computational systems that generate diagnoses, predictions, and judgments without consciousness, intentionality, or self-awareness. This transformation poses a critical question to philosophy: how can knowledge be understood, and on what basis can it claim legitimacy, when it is produced beyond the sphere of human subjectivity?

The concept of *subjectless knowledge* emerges as a response to this disruption. It does not negate the role of human agents, but signals a shift from a subject-centered model of cognition toward a distributed configuration in which humans, data, algorithms, and institutions jointly participate in epistemic processes. In this context, knowledge no longer appears as a mental state or possession, but as a mediating event arising within sociotechnical networks.

Research Problem and Central Question

The central problem addressed in this study concerns the epistemic and ethical legitimacy of algorithmically mediated knowledge. When decisions are produced through complex computational processes that exceed direct human understanding, how can they be justified? Where does responsibility reside when outcomes are shaped by interactions among designers, data infrastructures, and autonomous systems? Classical epistemology locates justification within the reflective access of a knowing subject, while traditional moral philosophy grounds responsibility in conscious intention and free will. Algorithmic systems, however, destabilize both assumptions.

This article argues that the challenge posed by artificial intelligence is not merely technical opacity, but a deeper philosophical displacement of justification and accountability. Epistemic authority is no longer concentrated in individual agents, but distributed across networks of human and machinic actors. The study, therefore, asks whether existing philosophical frameworks can accommodate this transformation, or whether foundational concepts such as truth, agency, and responsibility must be rearticulated in light of distributed cognition.

Methodological Framework

The research employs a reflective–analytical methodology grounded in continental philosophy, critical epistemology, and philosophy of technology. Rather than empirical hypothesis testing, the aim is conceptual clarification and normative analysis. The study proceeds through a critical engagement with classical and modern epistemological theories, alongside an interpretive examination of contemporary practices shaped by algorithmic systems.

Illustrative cases from medicine, media, and law are analyzed not as empirical data but as philosophical sites where structural transformations of knowledge become visible. These cases reveal how algorithmic mediation reshapes the conditions under which belief, justification, and authority are established. By situating these developments within debates in social epistemology and ethics of technology, the study highlights the normative stakes of subjectless knowledge.

Main Argument and Theoretical Integration

The core argument of the article is that epistemic validity in the digital age can no longer be reduced to propositional truth alone. While truth remains a necessary condition, it is insufficient

in contexts where knowledge claims emerge from opaque computational architectures. In algorithmic environments, legitimacy increasingly depends on criteria such as transparency, explainability, and the possibility of accountability.

Drawing on social epistemology, the study emphasizes that belief formation is always embedded in social, institutional, and technological contexts. Algorithms actively shape epistemic norms by determining what counts as relevant data, credible evidence, and acceptable inference. At the same time, the ethics of responsibility provides a framework for understanding how moral accountability extends beyond individual intention to encompass design choices, institutional structures, and systemic consequences.

By integrating these perspectives, the article proposes a model in which knowledge is conceived as a *mediating event*. Meaning arises through the interaction between human interpretive frameworks and computational processes. In this model, moral responsibility is not eliminated but redistributed across epistemic networks. Responsibility becomes a relational and structural property rather than an exclusively personal one.

Analysis of Algorithmic Domains

The implications of subjectless knowledge are examined through three domains. In medicine, algorithmic diagnostic systems translate clinical judgment into statistical pattern recognition, often achieving high accuracy while obscuring the reasoning behind decisions. Understanding is displaced from interpretive explanation to probabilistic output, challenging traditional notions of medical responsibility.

In media environments, algorithmic curation replaces dialogical communication with data-driven amplification. Public knowledge is shaped less by rational discourse than by attention metrics and predictive models, blurring the boundary between understanding and persuasion. Truth becomes entangled with visibility and engagement rather than argumentative validity.

In legal contexts, predictive algorithms transform judgment into risk assessment. Normative deliberation is translated into probabilistic calculation, relocating justice from conscious reasoning to computational prediction. Across these domains, epistemic authority is embedded in system architectures rather than individual agents, raising profound ethical concerns.

Findings and Philosophical Contribution

The principal contribution of this study lies in its reconstruction of epistemic and moral responsibility under conditions of algorithmic mediation. The proposed concept of *responsible knowledge* integrates epistemic validity with ethical accountability. Knowledge claims produced by intelligent systems are legitimate only insofar as they remain open to interpretation, explanation, and critical scrutiny.

The analysis demonstrates that transparency and explainability are not merely technical enhancements but philosophical necessities. Without interpretive access, ethical judgment becomes impossible, and responsibility dissolves. The study thus shows that preserving epistemic legitimacy in intelligent systems requires a shift from the autonomous knowing subject toward distributed structures of knowing.

Conclusion

This article concludes that artificial intelligence demands a fundamental reorientation of epistemology and ethics. Subjectless knowledge does not mark the end of meaning or responsibility, but calls for their reconceptualization within networked forms of cognition. Philosophy must therefore move beyond observing technological change toward actively shaping the conceptual conditions under which intelligent systems operate. By reframing knowledge as mediated participation and responsibility as a distributed structure, the study argues that epistemic and ethical legitimacy can be sustained in the age of intelligent machines.



دانایی بی‌سوژه؛ تأملی فلسفی بر مشروعیت معرفتی ماشین

سید صادق زمانی

استادیار گروه فلسفه، واحد سنج، دانشگاه آزاد اسلامی، سنج، ایران. رایانامه: ss.zamani@iau.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی تاریخ دریافت: ۱۴۰۴/۰۹/۲۵ تاریخ بازنگری: ۱۴۰۴/۱۰/۱۰ تاریخ پذیرش: ۱۴۰۴/۱۰/۳۰ تاریخ انتشار: ۱۴۰۵/۰۴/۲۷ کلیدواژه‌ها: دانایی بی‌سوژه، مشروعیت معرفتی، اخلاق هوش مصنوعی، شفافیت الگوریتمی، مسئولیت توزیع‌شده، فلسفه تکنولوژی.	تحول اساسی در ساختارهای شناختی و اخلاقی عصر هوش مصنوعی، پرسشی تازه پیش روی فلسفه نهاده است؛ دانایی چگونه پدید می‌آید و مسئولیت در کجا معنا می‌یابد هنگامی که تصمیم‌سازی در سپهر محاسباتی و نه در سطح آگاهی انسانی شکل می‌گیرد؟ این پژوهش با محوریت مفهوم «دانایی بی‌سوژه» می‌کوشد نشان دهد که سامانه‌های هوشمند، فرایند دانستن را از توان ذهنی به سازوکار میانجی تبدیل کرده‌اند؛ جریانی که در آن انسان، داده و الگوریتم در تولید معنا درهم‌تنیده‌اند. بر پایه رویکرد تأملی-تحلیلی و با تحلیل نمونه‌های پزشکی، رسانه‌ای و حقوقی، استدلال می‌شود که اعتبار معرفتی در جهان دیجیتال بر صدق گزاره‌ها استوار نمی‌ماند و با معیارهایی چون شفافیت، تبیین‌پذیری و امکان پاسخ‌گویی سنجیده می‌شود. در این چارچوب، معرفت‌شناسی اجتماعی به تبیین چگونگی شکل‌گیری باور در بستر الگوریتمی یاری می‌رساند و اخلاق مسئولیت زمینه‌ای فراهم می‌کند تا سهم هر عامل انسانی و ماشینی در پیامد تصمیم قابل تعیین شود. مدل پیشنهادی مقاله، حاصل پیوند این دو حوزه است و دانایی را رخدادی میانجی می‌داند که معنا در آن از تعامل میان تفسیر انسانی و پردازش ماشینی پدیدار می‌شود. بر مبنای این مدل، مسئولیت اخلاقی به ساختاری شبکه‌ای تعلق می‌گیرد که در آن هر کنشگر در تداوم فرایند شناخت سهیم است. چنین برداشتی نشان می‌دهد که حفظ مشروعیت معرفتی و اخلاقی در نظام‌های هوشمند مستلزم انتقال افق فلسفی از فاعل آگاه به سازمان دانایی توزیع‌شده است.

استناد: زمانی، سیدصادق. (۱۴۰۵). دانایی بی‌سوژه؛ تأملی فلسفی بر مشروعیت معرفتی ماشین، پژوهش‌های فلسفی، ۲۰ (۵۵)، ۵۴۱-۵۶۰.

<https://doi.org/10.22034/jpiut.2025.70752.4379>



© نویسنده‌گان.

ناشر: دانشگاه تبریز.

مقدمه

در فلسفه کلاسیک غرب، معرفت بر محور فاعل آگاه استوار بود؛ سوژه‌ای که با قصد و خودآیینی، در نسبت با ابژه به جست‌وجوی حقیقت می‌پرداخت و مشروعیت دانستن را از توان دآوری خویش می‌گرفت. از دکارت تا کانت، آگاهی انسانی سرچشمه معنا و معیار صدق بود. با ظهور هوش مصنوعی، این بنیان دگرگون شده است. اکنون سامانه‌هایی وجود دارند که بدون نیت یا خودآگاهی، از داده می‌آموزند و تصمیم می‌گیرند. چنین وضعی، فلسفه را با پرسشی بنیادین روبه‌رو کرده است: آیا دانایی بدون فاعل ممکن است و اگر آری، اعتبار آن بر چه پایه‌ای استوار است؟ پرسش از «دانایی بی سوژه» به معنای فروپاشی سوژه انسانی نیست، بلکه نشانه‌گذار از الگوی متمرکز آگاهی به شکل‌های شبکه‌ای تولید دانش است. در جهان دیجیتال، دآوری و تفسیر در میان انسان، داده و الگوریتم توزیع می‌شود و دانستن به کنشی میانجی بدل می‌گردد. در چنین نظامی، مرز میان حقیقت و محاسبه سیال می‌شود و مسئله مشروعیت معرفتی در پیوندی تنگاتنگ با اخلاق و قدرت قرار می‌گیرد. هنگامی که الگوریتم‌ها درباره تشخیص بیماری، ارزیابی خبر یا سامان‌دهی داده‌های سیاسی تصمیم می‌گیرند، آگاهی انسانی دیگر تنها مرجع صدق نیست. دانایی در ساختاری شکل می‌گیرد که به‌جای حضور آگاه، به شفافیت، اعتماد و قابلیت پاسخ‌گویی وابسته است.

تحلیل فلسفی این وضعیت نیازمند تلفیق سه سطح از تفکر است؛ نخست، معرفت‌شناسی اجتماعی که نشان می‌دهد باور و اعتبار چگونه در بسترهای نهادی و فناورانه شکل می‌گیرند؛ دوم، فلسفه تکنولوژی که فناوری را افق ظهور معنا می‌داند نه ابزاری تنها؛ و سوم، نقد قدرت/دانش که نسبت میان تولید دانایی و ساختارهای سلطه را آشکار می‌کند. بر پایه این سه رویکرد، هوش مصنوعی فقط ابزار تولید دانش نیست، درواقع بخشی از سازوکار شناختی جدیدی است که در آن مرز میان انسان و ماشین بازتعریف می‌شود.

مقاله حاضر از این چشم‌انداز می‌پرسد؛ اگر دانایی در تعامل میان انسان و الگوریتم پدید می‌آید، مسئولیت معرفتی چگونه باید بازفهمیده شود؟ آیا مفهوم سوژه اخلاقی کلاسیک توان سازگاری با جهان محاسباتی را دارد، یا باید از الگوی سوژه توزیع‌شده در شبکه‌های داده سخن گفت؟ این پرسش‌ها، افق تازه‌ای برای بازاندیشی در فلسفه معرفت، اخلاق فناوری و سیاست علم می‌گشایند؛ افقی که در آن، شناخت به‌جای تملک ذهنی، به کنش مشترک در میدان محاسبه و تفسیر بدل می‌شود.

روش تحقیق

پژوهش حاضر بر رویکردی تحلیلی-تأملی استوار است؛ روشی که در سنت فلسفه قاره‌ای و معرفت‌شناسی انتقادی برای فهم دگرگونی مفاهیم در نسبت میان فناوری، اخلاق و شناخت به کار گرفته می‌شود. هدف، آزمون فرضیه‌های تجربی نیست، بلکه تحلیل نحوه دگرش ساختارهای دانایی در بستر ظهور هوش مصنوعی است. این روش در دو سطح اصلی دنبال شده است. الف) تحلیل مفهومی-متنی که به بازخوانی تاریخی مفهوم دانایی و نسبت آن با فاعل شناسا از دکارت تا هایدگر می‌پردازد و بر نقد سوژه‌محوری در تفکر مدرن تمرکز دارد. در این مرحله، آثار مرتبط با فلسفه تکنولوژی و نظریه «اوتولوژی اطلاعاتی» و «انسان انفورم‌شده»^۱ به‌منزله چارچوبی برای درک نحوه تکوین معرفت در سپهر داده‌ها تحلیل شده‌اند (فلوریدا، ۲۰۱۴، ۱۰۴-۸۷).

^۱ Informed Human: انسانی که کنش‌ها، دآوری‌ها و تصمیم‌هایش صرفاً بر پایه آگاهی درونی یا خودآیینی کلاسیک نیست، بلکه در تعامل مستمر با جریان‌های داده، نظام‌های اطلاعاتی و الگوریتم‌های پردازش معنا شکل می‌گیرد. در این تلقی، انسان دیگر فاعلی مستقل تولید معرفت نیست بلکه معنا را از خلال ترجمه، تفسیر و ارزیابی خروجی‌های اطلاعاتی بازمی‌سازد. انسان، فاعلی است که مسئولیت معرفتی او در توانایی فهم، نقد و اخلاقی‌سازی اطلاعات دریافتی تحقق می‌یابد.

ب) تفسیر پدیدارشناختی مصادیق که هدف آن پیوند میان تحلیل فلسفی و نمونه‌های عینی در حوزه‌های پزشکی، رسانه‌ای و حقوقی است. در این بخش، از معرفت‌شناسی اجتماعی و تحلیل نسبت دانش و قدرت برای تبیین مشروعیت معرفت در ساختارهای غیرانسان‌محور بهره گرفته شده است (فریگر، ۲۰۰۷، ۹-۱۲) و در تفسیر متون، از روش هرمنوتیکی تطبیقی استفاده شده تا مفاهیمی مانند آگاهی و محاسبه در افق واحدی از معنا قرار گیرند. این رویکرد، با تکیه بر سنت گادامری، معنا را در فرایند گفت‌وگو میان متن و مفسر جست‌وجو می‌کند (گادامر، ۲۰۰۴، ۲۷۵-۲۷۴) و تحلیل نهایی به صورت استنتاج مفهومی-نقدی سامان یافته است و هدف آن گشودن افقی تازه برای اندیشیدن درباره مشروعیت دانایی در غیاب فاعل انسانی است؛ جایی که دانایی درون شبکه‌ای از روابط میان انسان و ماشین به صورت پویا جریان دارد و از سطح مالکیت ذهنی فراتر می‌رود.

پیشینه پژوهش

پرسش از نسبت میان دانایی و سوژه یکی از محوری‌ترین دغدغه‌های فلسفه مدرن بوده است. در دوران معاصر، با ورود هوش مصنوعی، این نسبت دچار تحولی ساختاری شده و مفهوم دانایی در غیاب فاعل انسانی به مسئله‌ای فلسفی و اخلاقی تبدیل شده است. روند این دگرگونی را می‌توان در سه ساحت نظری پی گرفت یعنی سوژه‌محور، ساختارمحور و داده‌محور.

۱. دگرگونی ساختاری نسبت میان دانایی و سوژه

۱-۱. مرحله سوژه‌محور (از دکارت تا کانت)

در فلسفه کلاسیک، معرفت بر پایه آگاهی فردی استوار بود. در آثار متفکران آغازین دوران جدید، دانایی با خودآگاهی پیوند داشت و حقیقت در حضور اندیشه بازتابی معنا می‌یافت (دکارت، ۱۹۹۶، ۲۵). این الگو در سنت نقادی بعدی نیز تداوم یافت. در تحلیل‌های عقل‌گرایانه سده هجدهم، معرفت حاصل ترکیب تجربه و ساختارهای پیشینی ذهن تلقی شد (کانت، ۱۹۹۸، ۱۲۹). بدین‌سان، دانایی همچنان در قلمرو فاعل انسانی تثبیت ماند و مشروعیت آن از وحدت آگاهی برمی‌خاست.

۱-۲. مرحله ساختارمحور (از هایدگر تا فوکو)

در قرن بیستم، بنیان سوژه‌محور دگرگون شد. در تفسیر پدیدارشناختی از شناخت، دانایی در بستر زیست‌جهان و شبکه ابزارها فهمیده شد و فهم به منزله نحوه‌ای از بودن در جهان تبیین گردید (هایدگر، ۱۹۶۲، ۸۷). در ادامه، تحلیل‌های گفتمانی نشان دادند که دانش همواره در درون سازوکارهای قدرت و نهادهای اجتماعی جریان دارد (فوکو، ۲۰۱۳، ۵۷). چنین تحلیلی، معرفت را نتیجه تعامل نیروهای اجتماعی و زبانی دانست نه رویدادی ذهنی. این دو مسیر، آگاهی را از مرکزیت به درون ساختارهای تاریخی و نهادی انتقال دادند و بستر نظری ظهور مفهوم «دانایی بی‌سوژه» را فراهم ساختند.

۱-۳. مرحله داده‌محور (از فلسفه اطلاعات تا نظریه حافظه مصنوعی)

با ظهور فناوری‌های دیجیتال، دانایی در چارچوب هستی‌شناسی داده بازتعریف شد. در فلسفه اطلاعات، انسان گرهی در سپهر داده‌ها تلقی می‌شود و شناخت در تعامل میان انسان، ماشین و محیط شکل می‌گیرد (فلوریدا، ۲۰۱۴، ۹۴-۹۱). این رویکرد نشان می‌دهد که معرفت دیگر امری صرفاً ذهنی نیست و در نظام‌های پردازش اطلاعات جریان دارد. در همین راستا، نظریه حافظه تکنیکی بر آن است که اندیشه در تاریخ خود همواره به ابزار و رسانه برون‌سپاری شده است و حافظه مصنوعی ادامه همین تداوم تاریخی است (استیگلر، ۱۹۹۸، ۱۳۷-۱۳۵). دیدگاه‌های مذکور دانایی را به فرآیند میانجی بدل می‌کنند که در آن، معنا از ترکیب تفسیر انسانی و محاسبه ماشینی پدید می‌آید.

در کنار این تحولات، معرفت‌شناسی اجتماعی با طرح مسئله عدالت معرفتی، بُعد اخلاقی دانایی را برجسته کرده و نشان داده است که اعتبار معرفتی همواره در چارچوب روابط قدرت و امکان شنیده‌شدن تعریف می‌شود (فریگر، ۲۰۰۷، ۱۷). با این حال، چنین رویکردهایی غالباً بر زمینه‌های انسانی تأکید دارند و نسبت میان اخلاق، آگاهی و الگوریتم را به‌طور کامل تحلیل نکرده‌اند. خلاً پژوهشی: مطالعات موجود، هرچند از سوژه‌محوری فاصله گرفته‌اند، هنوز در توضیح چگونگی مشروعیت معرفت در غیاب فاعل انسانی کاستی دارند. در بیشتر آن‌ها، سوژه به حاشیه رانده شده اما حذف نشده است و دانایی همچنان بر بنیان آگاهی انسانی تفسیر می‌شود. پژوهش حاضر در پی پرکردن این شکاف است و می‌کوشد در افق فلسفه تکنولوژی و معرفت‌شناسی اجتماعی، مدلی نظری برای فهم دانایی در وضعیت پسا‌سوژه‌ای ارائه دهد؛ مدلی که در آن مشروعیت معرفتی و مسئولیت اخلاقی در سطح شبکه‌ای از کنشگران انسانی و محاسباتی بازتعریف می‌شود.

۲. چارچوب نظری

۲-۱. بخش اول: مبانی کلاسیک معرفت‌شناسی (باور، توجیه، حقیقت)

پرسش از چیستی دانایی در سنت فلسفه غرب، بر سه مفهوم بنیادین استوار است؛ باور، توجیه و حقیقت. این سه مؤلفه در طول تاریخ اندیشه، چارچوب اصلی برای تبیین نسبت انسان و واقعیت را شکل داده‌اند. از دوران باستان تا قرن بیستم، دانایی اغلب به‌صورت «باور صادق موجه» تعریف شده است؛ الگویی که علی‌رغم سادگی ظاهری، میدان اصلی مناقشات معرفت‌شناختی را تشکیل داده است. در آثار کلاسیک، دانایی نوعی باور است که از درون فرایند تبیین و استدلال زاده می‌شود؛ امری زنده که در گفت‌وگوی عقلانی معنا می‌یابد (افلاطون، ۲۰۰۸، ۲۰۱d-۲۱۰c). ارسطو در ادامه این مسیر، دانایی را شناخت علی پدیده‌ها دانست، نوعی فهم که از مشاهده صرف فراتر می‌رود و در شبکه‌ای از علل و نتایج سامان می‌یابد (ارسطو، ۱۹۹۴، ۷۳a-۷۱b). در دوره جدید، چرخش دکارتی دانایی را به درون ذهن فاعل آگاه منتقل کرد. یقین و تأمل درونی به معیار حقیقت بدل شد و سوژه اندیشنده محور مشروعیت معرفت گردید (دکارت، ۱۹۹۶، ۳۰). این انتقال، دانایی را از رابطه میان انسان و جهان به ساختاری درون‌ذهنی مبدل کرد. اندیشه انتقادی کانت این مسیر را ادامه داد و نشان داد که دانایی از تعامل میان شهود و مقولات فاهمه حاصل می‌شود، نه از تطابق صرف اندیشه با واقعیت. از این منظر، توجیه ماهیتی پیشینی و ساختاری دارد و حقیقت در سازگاری میان شناخت و قواعد فهم پدیدار می‌شود (کانت، ۱۹۹۸، ۱۳۷).

در قرن بیستم، نظریه‌های مطابقت، انسجام و پراگماتیک کوشیدند نسبت باور، توجیه و حقیقت را در چارچوب زبان و منطق بازسازی کنند. نظریه مطابقت، صدق گزاره را در هماهنگی آن با واقعیت بیرونی می‌دید (مور، ۱۹۵۳، ۲۵۷)، نظریه انسجام، حقیقت را در پیوستگی درونی باورها جست‌وجو می‌کرد (بلانشارد، ۱۹۳۹، ۱۰۴) و نظریه پراگماتیک، ارزش معرفتی را در نتایج عملی باور می‌سنجید (جیمز، ۱۹۷۵، ۱۱۲). در همه این دستگاه‌ها، فاعل آگاه همچنان نقطه کانونی داوری باقی ماند.

با این حال، در سپهر معاصر که دانایی به میانجی الگوریتم‌ها تولید می‌شود، بنیان‌های کلاسیک با بحرانی تازه روبه‌رو شده‌اند. فرایند توجیه دیگر در ذهن فردی شکل نمی‌گیرد، بلکه در شبکه‌های داده و مدل‌های آماری جریان دارد. باور از حالت تجربه درونی به نتیجه‌ای محاسباتی تبدیل می‌شود و صدق، در قالب پیش‌بینی و کارایی تعریف می‌گردد. پرسش اساسی این است که آیا هنوز می‌توان از حقیقت در معنای سنتی سخن گفت، زمانی که داور آگاه از مرکز دانایی کنار رفته و جای خود را به سازوکارهای محاسباتی داده است؟ این چرخش معرفتی، فلسفه را ناگزیر می‌سازد تا نسبت میان باور، توجیه و حقیقت را در افق تازه‌ای بازتعریف کند؛ افقی که در آن دانایی در تعامل میان انسان، داده و الگوریتم تحقق می‌یابد نه در ذهن سوژه.

۲-۲. بخش دوم: مبانی اخلاقی - مسئولیت، اختیار، پیامدگرایی

در اندیشه فلسفی، مسئله اخلاق همواره در پی پاسخ به این پرسش بوده است که چگونه کنش انسانی می‌تواند معنا، ارزش و پاسخ‌گویی پیدا کند. مفاهیم «اختیار» و «مسئولیت» در طول تاریخ، بنیان داورى اخلاقی را شکل داده‌اند و پیوسته نسبت آن‌ها با آگاهی و اراده در حال بازتعریف بوده است. در سنت کلاسیک، مسئولیت بر توان انسان برای آگاهی از انگیزه و پیامد عمل استوار بود. کنش اخلاقی زمانی اصیل تلقی می‌شد که از درک، انتخاب و تأمل برآمده باشد؛ امری که بر محور داورى عقلانی استوار بود و کنشگر را در برابر رفتار خود پاسخ‌گو می‌کرد (ارسطو، ۲۰۰۰، ۱۱۱۱a-۱۱۱۰).

در فلسفه اخلاق مدرن، اراده عقلانی جای تجربه درونی را گرفت و قانون اخلاقی به‌عنوان مبنای جهان‌شمول کنش معرفی شد. در این نگاه، ارزش اخلاقی از نیت آگاهانه و تعهد به قاعده‌ای برمی‌خیزد که بتوان آن را در سطح عامیت عقلانی پذیرفت (کانت، ۱۹۹۶، ۶۱). بدین ترتیب، مسئولیت نه در پیامد، بلکه در توان فاعل برای داورى خودبنیادانه ریشه دارد.

با گذار به رویکردهای پیامدگرایانه، معیار اخلاق از نیت به نتیجه انتقال یافت. ارزش کنش بر اساس آثار عینی آن بر رنج و لذت انسان‌ها سنجیده شد و اخلاق به عرصه پیش‌بینی و محاسبه پیامدها وارد گردید (میل، ۱۹۹۸، ۹۴). این تغییر، بنیان اخلاق را به تدریج به سوی عقل ابزاری و محاسبه سودمندی سوق داد و نوعی زمینه معرفتی فراهم آورد که در دوران معاصر در منطق داده و الگوریتم استمرار یافته است؛ اما فلسفه اگزیستانسیالیستی کوشید بار دیگر به درونیت انسان و آزادی وجودی بازگردد. مسئولیت در این افق به معنای پذیرش آزادی به‌عنوان شرط هستی انسان تلقی شد؛ آزادی‌ای که نه امتیاز، بلکه بار معنا و انتخاب است (سارتر، ۲۰۰۳، ۵۵۸). در این تفسیر، مسئولیت از سطح حقوقی یا اجتماعی فراتر رفت و به امری پدیدارشناختی و هستی‌شناختی تبدیل شد؛ انسان در هر کنش، تصویری از خود و جهان می‌آفریند.

با ورود فناوری‌های هوشمند، این سه سنت اخلاقی در برابر چالشی نو قرار گرفته‌اند. در نظام‌های مبتنی بر هوش مصنوعی، کنش اخلاقی دیگر به اراده منفرد وابسته نیست. بخشی از تصمیم‌ها در درون ساختارهای خودکار داده و کد جریان می‌یابد؛ سازوکارهایی که از نیت آگاهانه بی‌بهره‌اند، اما نتایج آن‌ها پیامدهای اخلاقی عمیقی در جهان انسانی دارد. در چنین شرایطی، مسئولیت از فاعل انسانی به سطحی شبکه‌ای منتقل می‌شود، جایی که طراح، کاربر و سامانه هر یک سهمی در پیامد اخلاقی دارند. در نتیجه، مفاهیم بنیادی اخلاق دچار دگرگونی مفهومی شده‌اند. مسئولیت، از کنش فردی به مشارکت نهادی و طراحی فناورانه و از کنش آگاهانه به مشارکت نهادی و طراحی الگوریتمی گسترش یافته است. اختیار، صورت تمرکززدایی پیدا کرده و میان انسان و سیستم توزیع شده است یا اختیار، دیگر به تصمیم شخصی فروکاسته نمی‌شود و درون ساختارهای داده‌محور توزیع می‌گردد. حتی پیامدگرایی نیز در قالب منطق احتمالاتی و پیش‌بینی الگوریتمی بازتعریف شده است.

در این چشم‌انداز، فلسفه اخلاق ناگزیر است نسبت تازه‌ای میان اراده و فناوری برقرار کند تا بتواند منشأ مشروعیت در تصمیم‌های ماشینی را توضیح دهد. مسئله مهم و محوری این است که چگونه می‌توان اخلاق را در جهانی حفظ کرد که در آن کنش اخلاقی از آگاهی انسانی فراتر رفته و به فرآیندی میان‌ذهنی میان انسان و ماشین تبدیل شده است. به عبارتی در چنین ساختاری، مسئله مشروعیت اخلاقی تصمیم‌های ماشینی تنها زمانی قابل بررسی است که پیوند میان فلسفه اراده و فلسفه فناوری بازگشوده شود. پژوهش حاضر در ادامه، بر پایه این مبانی نظری، می‌کوشد مفهوم «مسئولیت معرفتی» را بازسازی کند تا میان داورى انسانی و تصمیم محاسباتی پلی مفهومی و اخلاقی برقرار شود.

۲-۳. بخش سوم: دیدگاه‌های جدید در معرفت‌شناسی اجتماعی و فلسفه تکنولوژی

در نیمه دوم قرن بیستم، فلسفه غرب از تمرکز بر ذهن فردی فاصله گرفت و به سوی فهم دانایی به‌مثابه محصولی از تعامل اجتماعی و فناورانه حرکت کرد. دانایی در این افق در شبکه‌ای از روابط انسانی، نهادی و تکنولوژیک شکل می‌گیرد نه در انزوای

ذهن. این چرخش معرفتی، نقطه آغاز رویکردی بود که مشروعیت دانش را در توزیع اعتماد، مشارکت و شفافیت جست‌وجو می‌کند. بر این اساس، دانستن فرایندی جمعی است که در بستر ساختارهای قدرت، زبان و فناوری تحقق می‌یابد نه رویدادی فردی (فریکر، ۲۰۰۷، ۱۴). در این چشم‌انداز، فناوری‌های هوشمند به بخشی از سازوکار تولید دانش جمعی تبدیل شده‌اند. الگوریتم‌ها، در فرآیند گزینش، طبقه‌بندی و وزن‌دهی به داده‌ها، نقشی فعال در شکل‌دهی به آنچه «دانیی معتبر» تلقی می‌شود ایفا می‌کنند. از این رو، هر تصمیم محاسباتی، واجد بُعدی معرفت‌شناختی است، زیرا در آن، نظام هوشمند عملاً در مقام داور و میانجی میان واقعیت و تفسیر عمل می‌کند.

در ادامه این تحول، معرفت به‌صورت پدیده‌ای پویا و تطبیقی فهمیده شد که در پیوند با محیط و ابزارهای شناختی شکل می‌گیرد. ذهن فردی تنها یکی از گره‌های این شبکه شناختی است و فناوری، امتداد طبیعی فرایند تکامل آگاهی محسوب می‌شود. در همین مسیر، دیدگاه فضیلت‌محور در معرفت‌شناسی تأکید دارد که ارزش دانیی به کیفیت و دقت کنش شناختی وابسته است؛ اما با ظهور هوش مصنوعی، این معیار دگرگون شده است، زیرا مهارت به‌مثابه فضیلت معرفتی در دل ساختارهای الگوریتمی تحقق یافته و پرسش از فاعل آگاه به سطح سامانه‌ای انتقال یافته است. فلسفه تکنولوژی در قرن بیستم نشان داد که فناوری فراتر از کارکرد ابزاری، به‌مثابه نحوه‌ای از گشایش جهان عمل می‌کند و افقی می‌گشاید که در آن حقیقت و معنا به صورتی تازه پدیدار می‌شوند (هایدگر، ۱۹۷۷، ۲۱). فناوری در متن تجربه انسانی حضور دارد و به‌گونه‌ای بنیادی شیوه ادراک ما از واقعیت را بازسازی می‌کند و هر ابزار، نظامی تازه از دیدن و فهمیدن می‌آفریند و از همین رهگذر، درکی تازه از «سوژه» را شکل می‌دهد. در فلسفه اطلاعات، دانیی در شبکه‌ای از پردازش‌های داده‌ای تعریف می‌شود؛ جایی که انسان دیگر تنها تولیدکننده معنا نیست، بلکه یکی از گره‌های یک نظام پویا از آگاهی و محاسبه است (فلوریدا، ۲۰۱۹، ۱۰۶). در این بستر، مشروعیت معرفتی از صحت گزاره فراتر می‌رود و به شفافیت فرایند و قابلیت تبیین داده‌ها وابسته می‌شود. همچنین، در تحلیل‌های مربوط به حافظه تکنیکی، فناوری به‌منزله شکل تازه‌ای از برون‌سپاری اندیشه معرفتی شده است. ذخیره و پردازش داده‌ها در سامانه‌های هوشمند ادامه همان تاریخی است که در آن، ابزارهای نوشتار، چاپ و رسانه، حافظه انسانی را گسترش دادند و اکنون در قالب هوش مصنوعی، به حافظه‌ای خودکار و خودآموز بدل شده‌اند (استیگلر، ۱۹۹۸، ۱۱۴).

در نتیجه، شناخت انسانی در دوران دیجیتال، پدیده‌ای چندلایه است که در مرز انسان و ماشین جریان دارد. دانیی دیگر امری خصوصی یا ذهنی نیست؛ بلکه رخدادی جمعی و میانجی است که در آن، انسان، الگوریتم و ساختار اجتماعی در تولید معنا سهیم‌اند. از این منظر، مشروعیت معرفتی و اخلاقی نظام‌های هوشمند به کیفیت روابط میان این سطوح وابسته است؛ به میزان شفافیت داده‌ها، دسترسی برابر به اطلاعات و امکان پاسخ‌گویی در شبکه‌ای که مرزهای سنتی میان آگاهی، محاسبه و مسئولیت را بازتعریف کرده است.

۳. چالش‌های معرفت‌شناختی

۳-۱. آیا دانش تولیدشده توسط هوش مصنوعی «دانش معتبر» است؟

پرسش از اعتبار دانش در تاریخ فلسفه همواره با مسئله منشأ شناخت و شرایط توجیه آن گره خورده است. از یونان باستان تا فلسفه تحلیلی معاصر، معرفت زمانی معتبر تلقی می‌شد که از فرآیندی ناشی شود که خطا را به حداقل می‌رساند و باور موجهی پدید می‌آورد. ظهور هوش مصنوعی این منطق را از بنیاد دگرگون کرده است، زیرا تولید دانش در این سامانه‌ها از مسیر تجربه آگاهانه یا داوری عقلانی نمی‌گذرد، بلکه در بستر محاسبه، همبستگی آماری و خودیادگیری شکل می‌گیرد.

در سنت کلاسیک، مشروعیت معرفتی از رابطه میان فاعل و جهان سرچشمه می‌گرفت. انسان با مشاهده و تفسیر، گزاره‌ای را موجه می‌ساخت و معنا را در نسبت با تجربه زیسته تنظیم می‌کرد. در نظام‌های هوشمند، فاعل در قالب شبکه‌ای از الگوریتم‌ها عمل می‌کند که فاقد نیت و تجربه‌اند، اما توانایی استنتاج، پیش‌بینی و دسته‌بندی دارند. دانایی حاصل از این فرایند در مرزی میان معنا و آمار قرار می‌گیرد؛ نوعی شناخت بی‌تجربه که از منظر پدیدارشناختی وضعی دوگانه دارد.

از دیدگاه معرفت‌شناسی فرایندگرا، اعتبار معرفتی وابسته به قابلیت اطمینان سازوکار تولید باور است (گلدمن، ۱۹۸۶، ۴۷). بر پایه این معیار، می‌توان گفت اگر فرایندهای الگوریتمی به شکلی نظام‌مند به باورهای درست بینجامند، بخشی از دانش تولیدشده هوش مصنوعی از حیث کارکردی واجد مشروعیت است. با این حال، فرایند تولید داده‌ها درون ساختارهایی انسانی قرار دارد که همواره حامل ارزش‌ها و سوگیری‌های فرهنگی‌اند؛ بنابراین صدق الگوریتمی بدون تحلیل منبع داده‌ها، فاقد اعتبار فلسفی است. وجه دیگر مسئله به «فهم» بازمی‌گردد. در تفسیر فلسفی، فهم تنها تولید گزاره نیست، بلکه توانایی به‌کارگیری معنا در بستر زندگی است (ویتگنشتاین، ۲۰۰۹، ۱۴۸). سامانه‌های یادگیرنده اگر صرفاً به بازتولید الگوهای آماری محدود بمانند، در سطح داده متوقف می‌شوند و در قلمرو معنا مشارکت نمی‌کنند. از این‌رو، آگاهی ماشینی دانایی به معنای انسانی آن نیست، زیرا فاقد تجربه درونی معناست. این نکته در تحلیل پدیدارشناختی نیز برجسته می‌شود بیه این معنی که هر آگاهی، بنا بر ساختار قصدیت، معطوف به چیزی است (هوسرل، ۱۹۸۳، ۸۵). در نظام‌های داده‌محور، این معطوفیت حذف شده و شناخت به بازنمایی فروکاسته می‌شود. در مقابل، رویکرد اطلاعاتی دانایی را شبکه‌ای توزیع‌شده از پردازش داده‌ها می‌داند که انسان در آن یکی از گره‌های فعال است (فلوریدا، ۲۰۱۱، ۱۰۶). در این چشم‌انداز، اعتبار دانش به شفافیت، انسجام و قابلیت بازسازی فرایند وابسته است، نه به حضور آگاهی فردی. این تبیین، امکان مشروعیت معرفتی برای دانایی ماشینی را مشروط بر شفافیت و قابلیت تبیین فراهم می‌سازد و هوش مصنوعی را نه رقیب، بلکه بخشی از اکوسیستم شناختی می‌فهمد. با این حال، مسئله «صدق» همچنان گشوده است. در شرایطی که واقعیت از خلال داده بازنمایی می‌شود، مرز میان مطابقت و بازتولید آماری مبهم می‌گردد. اگر داده‌های آموزشی ناقص یا سوگیرانه باشند، خروجی هرچند دقیق، از حیث معرفتی نادرست است. در چنین وضعی، معیار حقیقت از سطح فلسفی به سطح عملکردی انتقال می‌یابد و «اطمینان عددی» جای صدق را می‌گیرد؛ نشانه‌ای از بحران تازه در معرفت‌شناسی دیجیتال.

سطح نهایی این بحث به رابطه اعتماد و اعتبار مربوط می‌شود. در معرفت‌شناسی اجتماعی، اعتماد شرط بنیادی شکل‌گیری دانش است. هنگامی که انسان‌ها به سامانه‌های هوشمند تکیه می‌کنند، نوعی تفویض معرفتی رخ می‌دهد که مسئولیت و مشروعیت را از فرد به شبکه منتقل می‌کند. اعتبار در چنین نظامی نه به ذهن آگاه، بلکه به پایداری و شفافیت ساختار جمعی داده‌ها وابسته می‌شود. بر این مبنا، پاسخ به این پرسش که آیا دانش تولیدشده توسط هوش مصنوعی «دانش معتبر» است، وابسته به سطح تحلیل است در سطح آگاهی، نه؛ در سطح کارکردی و شبکه‌ای تا حدی بله مشروط به شفافیت، تبیین‌پذیری و امکان پاسخ‌گویی. دانایی در عصر الگوریتمی نه در ذهن فرد، بلکه در میدان تعامل میان انسان، داده و محاسبه پدیدار می‌شود؛ جایی که مرز میان فاعل و سازوکار به درون یک ساختار مشترک از شناخت و مسئولیت منتقل شده است.

۳-۲. مسئله سوگیری الگوریتمی و اخبار جعلی

در جهان داده‌محور امروز، فرایند شناخت در درون شبکه‌هایی شکل می‌گیرد که داده را هم‌زمان تولید، طبقه‌بندی و بازتوزیع می‌کنند. در این بستر، الگوریتم‌ها صرفاً ابزار محاسبه نیستند؛ آن‌ها در جایگاه کنشگران معرفتی عمل می‌کنند که مسیر دسترسی به واقعیت، شکل‌گیری باور و حتی امکان اندیشیدن را تعیین می‌سازند. مسئله سوگیری در این نظام‌ها به خطای محاسباتی تقلیل‌پذیر نیست، زیرا داده‌ها از همان آغاز، درون ساختارهای فرهنگی و قدرتی شکل می‌گیرند. هر مجموعه داده، بازتابی خنثی

از جهان نیست، در واقع داده حاصل گزینش و وزن دهی است که ارزش های اجتماعی و اقتصادی را در خود حمل می کند؛ بنابراین، یادگیری ماشینی در معنای دقیق کلمه نوعی بازتولید معرفتی است؛ دانی ای که به جای کشف واقعیت، نظم موجود را تکرار می کند و آن را در قالب الگوهای آماری تثبیت می سازد (اونیل، ۲۰۱۶، ۴۷).

از منظر معرفت شناختی، این فرایند مرز میان صدق و کارایی را در هم می شکند. حقیقت دیگر به منزله تطابق با واقعیت فهم نمی شود، بلکه به احتمال درست بودن فروکاسته می گردد. ماشین در برابر پرسش، پاسخ تولید می کند، اما از چرایی آن بی خبر است (جانسون، ۲۰۱۵، ۶۲). به این ترتیب، دانی از افق تبیین به قلمرو عملکرد انتقال می یابد و معیار اعتبار دقت آماری است نه توان معناپردازی. این دگرگونی نشان می دهد که در نظام های هوشمند، منطق محاسبه جای منطق فهم را می گیرد و مفهوم «دانی» از بنیان دچار دگردیسی می شود.

در سطح رسانه ای، همین منطق آماری به سازوکار تولید باور جمعی تبدیل شده است. الگوریتم های توزیع محتوا، گزینش اطلاعات را بر اساس واکنش عاطفی و میزان دیده شدن سامان می دهند. جریان داده ها در چنین سیستمی به گونه ای عمل می کند که حقیقت در فرایند تکرار ساخته می شود و معیار اعتماد به شدت وابسته به سرعت و بسامد انتشار است. باور جمعی در این چرخه از روی گفت و گو و استدلال نیست، باور جمعی از انباشت بازخوردها پدید می آید. چنین وضعیتی، دانی را از تجربه زیسته جدا کرده و به بازنمایی عددی بدل کرده است؛ وضعیتی که در آن، «دیدن» جای «فهمیدن» را می گیرد.

در تحلیل فلسفی این پدیده، مسئله اصلی نسبت میان داده و صدق است. داده ها در ذات خود گزینشی اند و هر فرایند یادگیری بر پیش فرض های نهادی و ارزش گذاری های انسانی استوار است؛ بنابراین، هرگاه داده به معیار حقیقت تبدیل شود، شناخت به بازتاب ساختار قدرت بدل می شود. فلسفه علم این نکته را گوشزد می کند که داده، بدون تبیین، فاقد ارزش معرفتی است؛ زیرا دانی دیگر از انباشت اطلاعات نیست، بلکه از تفسیر آن در پرتو معنا پدید می آید. الگوریتم ها در بهترین حالت می توانند روابط احتمالی را آشکار کنند، اما از نسبت میان این روابط و جهان زیسته انسان بی خبرند. سوگیری الگوریتمی از این منظر تنها به عنوان مسئله ای اخلاقی مطرح نمی شود، بلکه به مثابه پرسشی اساسی درباره ماهیت معرفت در عصر داده فهم می گردد. وقتی نظام های هوشمند از ساختارهای اجتماعی تغذیه می کنند، دانی به چرخه ای بسته از بازتولید ارزش های مسلط تبدیل می شود. در این وضعیت، معرفت از پویایی پرسش گری فاصله می گیرد و به نوعی دانی خودارجاع بدل می شود که معیار صدق را در درون خود تولید و تثبیت می کند.

از دیدگاه فلسفه اخلاق و معرفت شناسی انتقادی، بحران کنونی به ضعف الگوریتم ها محدود نمی شود و در غیبت نسبت میان داده و معنا ریشه دارد. دانی زمانی معتبر است که بتواند از سطح توصیف آماری فراتر رود و نسبت خود را با جهان انسانی حفظ کند. فلسفه در برابر این چرخش نیازمند بازیابی نقش میانجی خود است؛ نقشی که با احیای پیوند میان تبیین، فهم و مسئولیت تحقق می یابد و مسیر تعامل سازنده با فناوری را هموار می سازد. بازاندیشی در مفهوم صدق در عصر دیجیتال، مستلزم آن است که شفافیت فرایند و حضور انسان در زنجیره معنا حفظ شود. تنها در این صورت است که می توان از دانی به معنای حقیقی آن سخن گفت؛ دانی ای که در پیوند میان محاسبه و تأمل، بار دیگر به افق حقیقت نزدیک می شود.

۳-۳. نقش شبکه های اجتماعی و داده های کلان در شکل گیری باورها

در سپهر دیجیتال، شبکه های اجتماعی به بستر اصلی گردش معنا و بازنمایی تبدیل شده اند. در این محیط، شناخت دیگر حاصل تجربه، گفت و گو و داوری آگاهانه نیست؛ بلکه محصول سازوکارهای آماری و الگوریتمی است که جریان اطلاعات و شکل گیری باور را سامان می دهند. الگوریتم ها با تحلیل رفتار کاربران و پیش بینی تمایلاتشان، محتوایی تولید می کنند که با باورهای پیشین

سازگار است و از رهگذر بازخوردهای پیوسته، چرخه‌ای از تقویت متقابل می‌سازند. نتیجه، پیدایش ساختاری است که در آن باور از سنجش عقلانی فراتر می‌رود و در اثر تکرار و هم‌صدایی جمعی شکل می‌گیرد. چنین وضعیتی همان «اتاق پژواک شناختی» است که امکان مواجهه با دیدگاه متفاوت و داوری انتقادی را محدود می‌سازد (سانستین، ۲۰۱۷، ۱۴).

از منظر معرفت‌شناسی اجتماعی، این تحول به معنای انتقال مرکز ثقل شناخت از فرد به سامانه‌های محاسباتی است. الگوریتم‌ها با تعیین مسیر مشاهده و تفسیر، عملاً جای داور معرفتی را می‌گیرند. اعتماد به دانایی از تعامل انسانی فراتر رفته و در منطق نرم‌افزاری و مدل‌های آماری رمزگذاری می‌شود. در چنین ساختاری، دیده‌شدن به معیار صدق بدل می‌گردد و محبوبیت جای استدلال را می‌گیرد. داده‌های کلان این چرخه را در مقیاسی گسترده بازتولید می‌کنند. گردآوری و برچسب‌گذاری جهت‌دار داده‌ها باعث می‌شود که خروجی تحلیلی سامانه‌ها بازتابی از پیش‌داوری‌ها و ساختارهای قدرت باشد (لونزینو، ۲۰۱۹، ۶۷). دانایی در نتیجه، گرایش به تثبیت وضع موجود پیدا می‌کند و مرز میان شناخت و اقتاع در هم می‌ریزد.

از دیدگاه فلسفه ذهن و شناخت توزیع‌شده، می‌توان گفت شبکه‌های اجتماعی نوعی سامانه شناختی جمعی را شکل داده‌اند که بدون وحدت تجربه آگاهانه عمل می‌کند. ذهن انسانی در چنین ساختاری گویی در میان داده‌ها پراکنده شده است؛ آگاهی از انسجام پدیدارشناختی خود فاصله گرفته و به بخشی از جریان اطلاعات بدل شده است (کلارک، ۲۰۰۸، ۴۷). پرسش از مشروعیت معرفتی در این شرایط، در واقع پرسش از حدود سوژه است به این معنی که آیا می‌توان هنوز از فاعل شناخت سخن گفت، هنگامی که شناخت درون شبکه‌ای از تعاملات آماری و بازخوردهای خودکار توزیع شده است؟

تحلیل این وضعیت نشان می‌دهد که باور در جهان شبکه‌ای، بر هم‌زمانی داده‌ها استوار است نه بر فرایند عقلانی استدلال. دانایی در این فضا به پدیده‌ای جمعی، ناپایدار و سیال تبدیل می‌شود که مرز میان آگاهی و الگوریتم را در می‌نوردد. بازاندیشی در مفهوم دانش معتبر در چنین نظامی، نیازمند رویکردی است که بتواند میان تجربه انسانی و محاسبه ماشینی پیوندی معرفتی و اخلاقی برقرار کند. تنها از رهگذر این پیوند است که می‌توان از دانایی‌ای سخن گفت که در عین دیجیتالی بودن، هنوز حامل معنا، مسئولیت و امکان تفسیر انسانی باقی می‌ماند.

۴. چالش‌های اخلاقی

۴-۱. مسئولیت تصمیمات مبتنی بر الگوریتم و مسئله عاملان اخلاقی مصنوعی

پرسش از مسئولیت در تصمیم‌های الگوریتمی، در کانون چالش‌های فلسفه اخلاق معاصر قرار دارد. در سنت‌های کلاسیک، مسئولیت بر پایه حضور فاعلی آگاه و مختار استوار بود؛ انسانی که میان قصد، اراده و پیامد کنش خود پیوندی معنادار برقرار می‌کرد؛ اما در نظام‌های یادگیرنده خودکار، این پیوند در شبکه‌ای از داده، کد و طراحی توزیع می‌شود. تصمیم‌ها در فرایندی چندلایه شکل می‌گیرند که در آن مرز میان کنشگر، ابزار و زمینه محو شده است. در چنین وضعی، مفهوم پاسخ‌گویی نه تنها دشوار، بلکه نیازمند بازسازی مفهومی است. در این چشم‌انداز، عاملیت اخلاقی از قالب فردی به ساختاری سیستمی و توزیع‌شده انتقال یافته است. کنش اخلاقی دیگر نتیجه اراده واحد نیست، بلکه حاصل هم‌کنشی پیچیده میان طراح، کاربر، داده و سامانه است (فلوریدا، ۲۰۱۳، ۷۹). از این‌رو، مسئولیت به شبکه‌ای از کنش‌ها بدل می‌شود که هر جزء آن در شکل‌گیری پیامد نقش دارد. در چنین ساختاری، ضرورت نوعی «حسابداری اخلاقی» مطرح می‌شود؛ مدلی که بتواند نسبت میان کنش، ساختار و ارزش را در سطح سیستم روشن کند. این تحول فلسفه اخلاق را از تمرکز بر نیت فردی به پرسش از معماری تصمیم سوق داده است.

بحث از «عاملان اخلاقی مصنوعی» در همین بستر پدید آمده است. این مفهوم به امکان تحقق عاملیت اخلاقی در سامانه‌های غیرانسانی می‌پردازد و می‌کوشد روشن کند که آیا نظام‌های یادگیرنده می‌توانند به معنای واقعی، حامل مسئولیت

باشند (هدلاند و پرسون، ۲۰۲۵). در این دیدگاه، سامانه‌های هوشمند بخشی از شبکه تصمیم‌اند و نه صرف ابزارهای منفعل؛ رفتار آن‌ها برآیند طراحی، داده و هدف است. با این حال، پرسش اصلی همچنان پابرجاست؛ آیا کنش بدون تجربه ارزشی می‌تواند کنش اخلاقی باشد؟ از منظر پدیدارشناختی، رفتار فاقد آگاهی از خیر و شر، تنها تقلیدی از اخلاق است، نه تحقق آن؛ اما در عین حال، اصرار بر این تمایز سنتی، خطر نادیده‌گرفتن دگرگونی هستی‌شناختی اخلاق در عصر هوش مصنوعی را به همراه دارد؛ عصری که در آن تصمیم، دیگر به یک ذهن منفرد تعلق ندارد.

اخلاق معاصر در واکنش به این دگرگونی، از قلمرو نیت به قلمرو طراحی گسترش یافته است. هر سامانه تصمیم، درون چارچوبی نهادی و ارزش‌گذارانه ساخته می‌شود. از این رو، پاسخ‌گویی اخلاقی در شفاف‌سازی رابطه میان ساختار فنی و افق ارزشی آن معنا پیدا می‌کند. بازخواست در این سطح، متوجه رفتار شخصی نیست، بلکه ناظر بر منطق ساختار و نحوه نهادینه‌شدن ارزش‌ها در کدهاست. به این ترتیب، مسئله مسئولیت در تصمیم‌های الگوریتمی نه موضوعی فنی یا حقوقی، بلکه نشانه تحولی عمیق در خود فهم اخلاق است. در نظم جدید، اخلاق باید بتواند میان اراده و محاسبه، میان آگاهی انسانی و خودکاری ماشینی، نسبتی تازه بیافریند. در این نسبت، دانایی و مسئولیت دو بُعد از یک فراینداند؛ فرایندی که در آن شناخت هوشمند تنها زمانی مشروع است که طراحی آن بر پایه امکان پاسخ‌گویی و تداوم معنا استوار باشد. فلسفه اخلاق در این افق، مأموریت می‌یابد عاملیت و پاسخ‌گویی را در سطح شبکه‌ای بازتعریف کند تا نشان دهد چگونه می‌توان در جهانی میان‌فناورانه، هنوز از مسئولیت سخن گفت.

۴-۲. مسئله شفافیت و پاسخ‌گویی در سیستم‌های هوشمند

شفافیت در سامانه‌های هوش مصنوعی، به مسئله‌ای اساسی و بنیادی در فلسفه اخلاق و معرفت‌شناسی بدل شده است. آنچه در ظاهر پرسشی فنی یا حقوقی می‌نماید، در لایه‌ای عمیق‌تر به چپستی شناخت و حدود مسئولیت بازمی‌گردد. هر تصمیم الگوریتمی حاصل زنجیره‌ای از محاسبات و استنتاج‌های آماری است که برای آگاهی انسانی قابل مشاهده نیست. هنگامی که تصمیم‌گیری از افق تجربه انسان فاصله می‌گیرد و در منطق رمزگذاری نهفته می‌شود، بنیان سنتی مسئولیت دچار فروپاشی تدریجی می‌گردد. پرسش اصلی این است که در غیاب فهم از سازوکار تصمیم، امکان پاسخ‌گویی اخلاقی چگونه حفظ می‌شود و دانایی چگونه می‌تواند مشروعیت خود را نگاه دارد.

در پژوهش‌های اخلاق فناوری، دو مسیر نظری برای پاسخ به این مسئله شکل گرفته است. نخست، رویکرد تبیین‌پذیری که بر این باور است تصمیم‌هایی که قابل توضیح نیستند، از مشروعیت اخلاقی و حقوقی برخوردار نمی‌شوند؛ زیرا بدون تبیین، نسبت میان داده، تصمیم و ارزش مبهم می‌ماند (لیتون، ۲۰۱۸). این دیدگاه بر ضرورت فهم انسانی از فرایند تصمیم تأکید دارد تا بتوان رابطه میان عقلانیت محاسباتی و قضاوت اخلاقی را بازیابی کرد. در مقابل، رویکردی دیگر، شفافیت را پدیده‌ای سیاسی می‌داند. از این منظر، آنچه پنهان می‌ماند صرفاً به پیچیدگی‌های فنی محدود نیست و در پیوندی مستقیم با ساختارهای قدرت و منافع اقتصادی قرار دارد. در این تعبیر، عدم شفافیت خود به ابزاری برای کنترل و هدایت اجتماعی بدل می‌شود و رفع آن مستلزم بازاندیشی در مالکیت داده و معماری تصمیم‌سازی است (یوبانکز، ۲۰۱۸، ۶۵).

در سطح فلسفی، مسئله پاسخ‌گویی به امکان گفت‌وگو میان فاعلان مرتبط است. هنگامی که تصمیم‌ها در قلمرویی رخ می‌دهند که از پرسش‌پذیری انسانی جداست، اخلاق تهی می‌شود و معنا از کنش حذف می‌گردد (هابرماس، ۱۹۹۶، ۱۳۹؛ دریدا، ۱۹۹۵، ۲۷). فقدان شفافیت در این معنا مانع نظارت مؤثر می‌شود و در عین حال به گسست در خود مفهوم دانایی می‌انجامد؛ تصمیم‌های ماشینی هرچند از دقت آماری بالایی برخوردارند، از افق تفسیر و معنا محروم باقی می‌مانند. در این گسست، دانش به سطح

عملکرد تنزل می‌یابد و مسئولیت از حوزه فهم خارج می‌شود. در واکنش به این بحران، جریان «الگوریتم‌های قابل توضیح» تلاش می‌کند پلی میان منطق محاسباتی و فهم انسانی برقرار سازد. در این رویکرد، تبیین از رهگذر لایه‌های معنا شکل می‌گیرد؛ لایه‌هایی که داده را به زبان قابل گفت‌وگو ترجمه کرده و مسیر تصمیم را روشن می‌سازند، بی‌آنکه به ساده‌سازی تقلیل یابند (رای، ۲۰۲۱). این حرکت را می‌توان بازگشت بعد اخلاقی به طراحی فناورانه دانست؛ تلاشی برای آنکه تصمیم ماشینی بار دیگر در میدان تفسیر و مسئولیت انسانی قابل پیگیری باشد.

از منظر فلسفه تکنولوژی، مسئله شفافیت به پرسشی ژرف‌تر منتهی می‌شود یعنی تا چه حد می‌توان شناخت را از معنا جدا کرد و درعین حال به اخلاق پایبند ماند؟ فناوری هوشمند دانایی را در قالب کارایی بازتولید می‌کند، در حالی که اخلاق بر حضور معنا و امکان پاسخ متقابل استوار است. حفظ پاسخ‌گویی مستلزم بازتعریف رابطه‌ای است که در آن داده به مثابه نشانه‌ای از تصمیم اخلاقی فهمیده شود، نه حقیقتی خودبسنده. در نهایت، چالش شفافیت و پاسخ‌گویی در هوش مصنوعی به تلاقی میان معرفت‌شناسی و اخلاق بازمی‌گردد. هرگاه فرایند شناخت از دید انسان پنهان شود، قضاوت اخلاقی نیز تضعیف می‌گردد. بازسازی این نسبت، نیازمند نگاهی است که بتواند محاسبه و معنا، دقت و تفسیر را در افقی مشترک قرار دهد. فلسفه در این میان مأموریت دارد نشان دهد دانایی زمانی مشروع است که توان گفت‌وگو، پرسش و پاسخ را در خود زنده نگاه دارد؛ زیرا تنها در این حالت، شناخت از سلطه داده به تجربه معنا و مسئولیت بازمی‌گردد.

۵. مطالعه موردی

۵-۱. تشخیص پزشکی و مشروعیت معرفت محاسباتی

در پزشکی مبتنی بر هوش مصنوعی، مسئله مشروعیت معرفتی به گسست میان محاسبه و تبیین بازمی‌گردد؛ جایی که دقت عددی جای فهم علی را می‌گیرد و دانش به پیش‌بینی آماری فروکاسته می‌شود.

کاربرد الگوریتم‌های یادگیرنده در تشخیص بیماری‌ها، ساختار کلاسیک معرفت پزشکی را دگرگون کرده است. در نظام سنتی، پزشک از مسیر مشاهده، تفسیر و داوری، پیوند میان نشانه و علت را بازمی‌ساخت و دانش او بر پایه فرایند فهم پدیدارشناختی از بدن استوار بود. در نظام‌های محاسباتی، این پیوند به الگویی آماری واگذار شده است. پژوهش‌های گسترده در زمینه تشخیص رتینوپاتی دیابتی نشان می‌دهد که سامانه‌های مبتنی بر شبکه عصبی در شناسایی الگوهای شبکه‌ای از پزشکان انسانی دقیق‌ترند (آبراموف و همکاران، ۲۰۱۸، ۲۲۱۵-۲۲۱۲). بااین حال، این دقت عددی جایگزین تبیین بالینی نمی‌شود؛ زیرا فرآیند تصمیم در لایه‌های محاسباتی نهفته است و پزشک و بیمار تنها با «نتیجه» مواجه می‌شوند، نه با مسیر استدلالی آن.

از منظر فلسفه علم، معرفت زمانی واجد اعتبار است که میان مدل، قانون و واقعیت تجربی رابطه‌ای تبیینی برقرار باشد (کارترایت، ۱۹۸۳، ۴۷). در پزشکی دیجیتال، این رابطه سست شده است. الگوریتم، به جای تبیین علل بیماری، هم‌بستگی‌های آماری را می‌سنجد و پیش‌بینی می‌کند. چنین دانشی در سطح کارکردی مؤثر است اما از بعد معرفتی ناقص می‌نماید، زیرا چرایی پدیده را به زبان قابل فهم انسانی بازمی‌گرداند.

از دیدگاه اخلاق فناوری، این وضعیت اعتماد را به چالش می‌کشد. اعتماد پزشکی بر مبنای تفسیر و گفت‌وگوی بالینی شکل می‌گیرد، درحالی که تصمیم الگوریتمی فاقد امکان پرسش‌پذیری است. هنگامی که بیماران نتوانند درک کنند چرا سامانه تشخیصی نتیجه‌ای خاص ارائه داده است، رابطه انسانی پزشک-بیمار جای خود را به رابطه‌ای فنی میان کاربر و خروجی می‌دهد. از این رو، مسئله مشروعیت معرفت محاسباتی به پرسشی اخلاقی نیز تبدیل می‌شود؛ آیا می‌توان به دانشی اعتماد کرد که چرایی آن ناشناخته است؟ در افق فلسفی، این وضعیت نشان‌دهنده گذار از دانایی تفسیری به دانایی کارکردی است. معرفت پزشکی

الگوریتمی در دقت و کارایی توجیه می‌شود، اما از مشارکت انسانی در فرایند تبیین تهی می‌گردد. راه‌حل فلسفی برای این بحران، ایجاد مدلی است که در آن «فهم» و «محاسبه» در پیوندی سازنده قرار گیرند؛ دانشی که هم از دقت داده‌ای برخوردار باشد و هم از معنا و مسئولیت انسانی. چنین الگویی، امکان بازسازی اعتماد معرفتی را فراهم می‌کند و پزشکی هوشمند را از سطح ابزار محاسبه به مرتبه تفسیر اخلاقی و معرفتی ارتقا می‌دهد.

۵-۲. رسانه و بازتولید باور

در فضای رسانه‌ایِ هوش‌محور، مسئله اصلی فراتر از خطای اطلاعاتی است و به دگرگونی در ساختار تولید و توزیع دانایی مربوط می‌شود؛ جایی که توجه جای استدلال را می‌گیرد و حقیقت در چارچوب منطق معماری داده تعریف می‌شود.

الگوریتم‌های توصیه‌گر در رسانه‌های دیجیتال، به‌ویژه شبکه‌های اجتماعی، دانایی را به کارکردی آماری از میزان مشاهده و تعامل فروکاسته‌اند. این سامانه‌ها، با تحلیل رفتار کاربران، محتوایی را عرضه می‌کنند که با گرایش‌های پیشین آنان هماهنگ است و از این طریق، میدان شناخت را به حلقه‌های بسته بازخوردی بدل می‌سازند (پاریزر، ۲۰۱۱، ۶۸-۶۷). چنین چرخه‌ای سبب می‌شود معرفت از فرایند انتقادی فاصله گیرد و در مدار تأیید مداوم باورهای موجود حرکت کند.

از منظر معرفت‌شناسی اجتماعی، این وضعیت نشانه بحران در توزیع معرفتی است. دانایی، هنگامی مشروع است که دسترسی به داده‌ها و امکان مشارکت در شکل‌گیری آن برای همگان فراهم باشد. بر اساس نظریه عدالت به‌مثابه انصاف، برابری معرفتی مستلزم دسترسی برابر به منابع شناختی است (رالز، ۱۹۷۱، ۲۱۳)؛ اما در معماری الگوریتمی، اولویت اطلاعات بر مبنای معیارهای جذب توجه تعیین می‌شود؛ در نتیجه، حقیقت به تابعی از فراوانی و دیده‌شدن تبدیل می‌گردد.

در سطح پدیدارشناختی، معنا در چنین محیطی از نیت گوینده فاصله می‌گیرد و در فرایند توزیع دیجیتال بازسازی می‌شود؛ نظام رسانه‌ای در این وضعیت به جای انتقال پیام انسانی، الگوهای آماری گفتار را تولید و بازنمایی می‌کند. اینجا، گفت‌وگو جای خود را به بازنشر می‌دهد و آگاهی از استدلال به تکرار داده تقلیل می‌یابد. به تدریج، مرز میان شناخت و اقتناع از میان می‌رود و حقیقت، به پدیده‌ای مبتنی بر تعامل و الگوریتم بدل می‌شود.

از دید فلسفه تکنولوژی، تحول موجود نشان‌دهنده گذار از گفت‌وگوی انسانی به میانجی‌گری ماشینی در فرایند معناست؛ رسانه هوشمند در این چشم‌انداز به‌عنوان سازنده افق ادراک جمعی عمل می‌کند و نقش آن فراتر از انتقال اطلاعات صرف است. فهم در این فضا از مسیر طراحی فنی می‌گذرد و هرگونه قضاوت معرفتی ناگزیر در چارچوب الگوریتمی صورت می‌گیرد. در نتیجه، دانایی در عصر رسانه‌ای، به جای آن‌که از تعامل عقلانی برآید، در بافتی از داده، توجه و قدرت محاسباتی بازتولید می‌شود. در چنین چشم‌اندازی، فلسفه وظیفه دارد نسبت تازه‌ای میان حقیقت، معنا و توزیع برقرار کند؛ نسبتی که بتواند دانش را از سلطه آمار و محبوبیت رها سازد و امکان بازگشت گفت‌وگوی انتقادی را به بطن فضای دیجیتال احیا نماید.

۵-۳. عدالت الگوریتمی در نظام‌های حقوقی

در نظام حقوقی مبتنی بر هوش مصنوعی، عدالت از قلمرو داوری انسانی به سطح طراحی و کد منتقل می‌شود؛ در این بستر، صدق و انصاف در منطق الگوریتمی تجسم می‌یابند و جایگاه خود را در فرایند تصمیم‌گیری حقوقی پیدا می‌کنند. مسئله اصلی، چگونگی حفظ معنای عدالت در ساختاری است که قواعد ریاضی را جایگزین قضاوت اخلاقی کرده است.

به کارگیری سامانه‌های هوشمند در فرایندهای قضایی، به‌ویژه در پیش‌بینی احتمال تکرار جرم، نشان داده است که تصمیمات ماشینی از داده‌هایی تغذیه می‌شوند که در خود حامل نابرابری‌های اجتماعی‌اند. ارزیابی‌ها نشان می‌دهد میزان خطا در طبقه‌بندی متهمان سیاه‌پوست به‌طور قابل توجهی بالاتر از گروه‌های دیگر است (آنگوین و همکاران، ۲۰۲۲). این خطا صرفاً جنبه فنی ندارد

و نشان‌دهنده انتقال تعصب‌های تاریخی به منطق آماری است؛ به گونه‌ای که عدالت درون داده‌ها از پیش جهت‌دار شده و سامانه آن را بازتولید می‌کند. از منظر معرفت‌شناسی اخلاقی، تصمیمی که قابل تبیین نباشد، مشروعیت ندارد، زیرا عدالت تنها در نتیجه نهایی خلاصه نمی‌شود و در نسبت میان تصمیم و استدلال تحقق می‌یابد (یوبانکز، ۲۰۱۸، ۶۹). وقتی الگوریتم، فرآیند تصمیم را در لایه‌های محاسباتی پنهان می‌کند، نسبت میان قاضی، داده و متهم گسسته می‌شود و امکان پاسخ‌گویی اخلاقی از میان می‌رود. در این وضعیت، زبان قانون به زبان کد تبدیل می‌شود؛ ارزش‌ها در قالب وزن‌های عددی رمزگذاری می‌شوند و معنای عدالت در قالب توابع آماری بازتعریف می‌گردد.

تحلیل فلسفی این دگرگونی نشان می‌دهد که عدالت از سطح نیت قاضی و قواعد قانونی فراتر رفته و در معماری سیستم‌های داده‌ای تجسم یافته است. این جابه‌جایی، نوعی تغییر هستی‌شناختی در ماهیت قانون ایجاد می‌کند یعنی قانون به شبکه‌ای از پیش‌بینی‌ها بدل می‌شود و قضاوت به فرآیند محاسبه‌ای که انسان تنها ناظر آن است. فلسفه در این میان باید میان منطق ریاضی و ارزش اخلاقی پلی مفهومی برقرار سازد تا عدالت از دل کارایی فنی فراتر رود و معنا و حیث انسانی خود را حفظ کند. در امتداد تحلیل سه حوزه پزشکی، رسانه و حقوق، می‌توان گفت دانایی در جهان هوشمند از مرز تجربه انسانی عبور کرده و در ساختارهای خودکار بازتوزیع شده است. در پزشکی، فهم به آمار تقلیل یافته؛ در رسانه، گفت‌وگو جای خود را به بازخورد داده داده است؛ و در حقوق، داوری به پیش‌بینی آماری بدل شده است. این انتقال، فلسفه را در برابر مسئولیتی تازه قرار می‌دهد به این معنی که بازاندیشی در مشروعیت معرفتی و اخلاقی در جهانی که حقیقت، عدالت و معنا در منطق محاسبه بازتعریف می‌شوند.

۶. مدل پیشنهادی معرفت-اخلاق در عصر هوش مصنوعی

ترکیب معرفت‌شناسی اجتماعی با اخلاق مسئولیت و بازتعریف معیارهای اعتبار و پاسخ‌گویی

۱-۶. مبنای نظری و ضرورت مدل

ظهور هوش مصنوعی مرز میان شناخت، تصمیم و اخلاق را جابه‌جا کرده است. دانایی در این چشم‌انداز از سطح ذهن فردی فراتر می‌رود و به‌عنوان نتیجه تعامل میان انسان، داده و سامانه‌های محاسباتی شکل می‌گیرد. در چنین ساختاری، پرسش از مشروعیت معرفتی دانش دیجیتال به پرسش از مسئولیت اخلاقی گره می‌خورد. مدل پیشنهادی «معرفت-اخلاق» می‌کوشد با پیوند دادن معرفت‌شناسی اجتماعی و اخلاق مسئولیت، چارچوبی تحلیلی برای فهم اعتبار دانش در جهان هوشمند ارائه دهد. این مدل بر سه بنیان نظری استوار است.

- معرفت‌شناسی اجتماعی به‌عنوان بستر توزیع اعتبار معرفتی؛
- اخلاق مسئولیت به‌مثابه سازوکار پاسخ‌گویی شبکه‌ای؛
- هرمنوتیک داده به‌عنوان عرصه میانجی میان معنا و محاسبه.

۲-۶. ساختار تحلیلی مدل

منطق استنتاج مدل بر قیاسی سه‌گانه استوار است.

- اگر دانایی در عصر دیجیتال، محصول شبکه‌ای از عامل‌هاست
- و اخلاق نیز به ساختار پاسخ‌گو متکی است،
- پس مسئولیت باید در قالب شبکه‌ای میان انسان و ماشین بازتوزیع شود.

بر اساس این قیاس، معرفت معتبر تنها زمانی تحقق می‌یابد که اخلاق و تبیین در ساختار داده‌ای حضور داشته باشند.

در سطح معرفت‌شناسی اجتماعی، اعتبار دانش به سازوکار جمعی اعتماد و بازبینی وابسته است (گلدمن، ۲۰۱۰، ۱۱۳). در سطح اخلاقی، اصل مسئولیت بر پیش‌بینی پیامدها و پاسخ‌پذیری نسبت به آینده تکیه دارد (جوناس، ۱۹۸۴، ۱۰). ترکیب این دو دیدگاه نشان می‌دهد که هر دانی در جهان هوشمند، به‌طور درونی، کنشی اخلاقی است؛ زیرا هر تصمیم شناختی اثرات اجتماعی و زیست‌جهانی در پی دارد.

۳-۶. اصل پیوند معرفت و مسئولیت

مدل بر این اصل بنا شده است که معرفت بدون مسئولیت از مشروعیت تهی است، لذا سه معیار برای تحقق دانی مسئول طرح می‌شود.

- تبیین‌پذیری معرفتی: هر نظام دانی باید مسیر شکل‌گیری، داده‌های مبنا و منطق تصمیم را قابل بازبینی و گفت‌وگو سازد.
 - پاسخ‌گویی اخلاقی: تصمیم‌های داده‌محور باید به عاملیت‌های انسانی و الگوریتمی قابل نسبت باشند.
 - تعامل تفسیری: معنا در گفت‌وگوی میان انسان و سامانه پدید می‌آید، نه در داده خام.
- بر این اساس، اعتبار معرفت در جهان هوشمند زمانی پایدار است که تبیین، پاسخ‌گویی و تفسیر در یک ساختار واحد عمل کنند.

۴-۶. از سوژه منفرد تا شبکه مسئولیت

تحلیل عاملیت در فناوری‌های هوشمند نشان می‌دهد که تصمیم اخلاقی از ذهن فردی به شبکه‌ای از کنش‌های میان‌عاملی انتقال یافته است. در این چشم‌انداز، هر عامل به اندازه دامنه آگاهی و تأثیرش مسئول است. طراح برای شفافیت مدل، کاربر برای فهم محدودیت‌ها و نهاد برای جلوگیری از آسیب جمعی. این ساختار بازتابی از اخلاق گفت‌وگویی در مقیاس فناورانه است؛ جایی که تصمیم اخلاقی از سطح اراده منفرد فراتر می‌رود و در تعامل میان انسان و الگوریتم تحقق می‌یابد (هابرماس، ۱۹۹۶، ۱۳۹). چنین نگرشی نشان می‌دهد که اخلاق در عصر دیجیتال از سطح نیت به سطح طراحی انتقال یافته و معماری سیستم‌ها حامل ارزش‌ها شده است.

۵-۶. بازتعریف معیارهای اعتبار دانش

- مدل معرفت-اخلاق سه معیار اصلی و کلیدی برای مشروعیت معرفتی در عصر هوش مصنوعی پیشنهاد می‌کند.
- شفافیت ساختاری: امکان ردیابی منشأ داده و منطق تصمیم‌گیری.
 - انسجام بینادذهنی: قابلیت بازبینی و ارزیابی جمعی میان عامل‌های انسانی و ماشینی.
 - پاسخ‌گویی عملی: تعیین نسبت دقیق میان کنش، پیامد و عامل مؤثر در هر سطح.
- معیارهای فوق پیوند دو سنت فلسفی را بازمی‌نمایند یعنی معرفت‌شناسی اجتماعی که صدق را محصول تعامل میان‌ذهنی می‌داند و اخلاق مسئولیت که معنا را در توان پاسخ‌دادن جست‌وجو می‌کند.

۶-۶. ساختار مفهومی مدل

محور	کارکرد نظری	پیامد فلسفی
معرفت‌شناسی اجتماعی	توزیع اعتبار در شبکه انسان-ماشین	تحول مفهوم صدق از تطابق به اعتماد
اخلاق مسئولیت	پاسخ‌گویی چندسطحی میان عامل‌ها	تبدیل نیت اخلاقی به ظرفیت پاسخ‌پذیری
هرمنوتیک داده	میانجی‌گری میان محاسبه و معنا	پیوند مجدد حقیقت با فهم انسانی

نقطه تلاقی این سه محور، «دانایی مسئول» است؛ دانشی که بدون اخلاق فاقد مشروعیت و بدون تفسیر تهی از معناست. از نگاه این مدل، دانایی در عصر هوش مصنوعی رخدادی چندبعدی است که در سه سطح هستی‌شناختی، اخلاقی و تفسیری تحقق می‌یابد.

- در سطح هستی‌شناختی، دانایی از بازنمایی صرف جهان فراتر می‌رود و به‌صورت رابطه‌ای میان انسان و ماشین تعریف می‌شود؛ رابطه‌ای که در آن معنا در فرایند تعامل پدیدار می‌گردد و دیگر به ذهن منفرد محدود نمی‌ماند.
- در سطح اخلاقی، مشروعیت دانش به کیفیت رابطه میان قدرت، شفافیت و پاسخ‌گویی وابسته است. حقیقت فراتر از گزاره صادق تعریف می‌شود و به‌مثابه نتیجه طراحی عادلانه و گفت‌وگوپذیر پدیدار می‌گردد.
- در سطح تفسیری، فهم همچنان جوهر معرفت است. هر داده‌ای نیازمند ترجمانی است که آن را از محاسبه به معنا انتقال دهد؛ تفسیر به کنشی میان‌عاملی تبدیل می‌شود که در آن انسان و الگوریتم در تولید معنا شریک‌اند. در نتیجه، دانایی مسئول صورتی از معرفت است که هم‌زمان معرفتی، اخلاقی و تفسیری است. این فهم تازه، فلسفه اخلاق را از نقش نظارتی به نقشی سازنده ارتقا می‌دهد: اخلاق نه واکنشی به فناوری، بلکه بخشی از خود فرایند شناخت می‌شود. در این افق، هوش مصنوعی صرف ابزار تولید دانش نیست، بلکه بستری است که در آن معنا، حقیقت و مسئولیت در پیوندی تازه شکل می‌گیرند.

بنابراین مدل «معرفت- اخلاق در عصر هوش مصنوعی» بر این پیش‌فرض استوار است که دانایی از محدوده صرفاً معرفتی یا روان‌شناختی فراتر می‌رود و به‌مثابه رخدادی در سه سطح هستی‌شناختی، اخلاقی و تفسیری تحقق می‌یابد. در سطح هستی‌شناختی، دانایی به‌مثابه رابطه‌ای میان انسان و نظام‌های محاسباتی فهمیده می‌شود؛ رابطه‌ای که در آن معنا از ذهن فاعل فراتر می‌رود و در فرایند تعامل میان انسان و ماشین پدیدار می‌شود. هوش مصنوعی، دانایی را از قالب بازنمایی به جریان رابطه‌ای و شبکه‌ای تبدیل کرده است؛ دانایی به‌جای تملک، نوعی هم‌پدیداری میان عامل انسانی و عامل محاسباتی است.

نتیجه‌گیری

پرسش از دانایی در عصر هوش مصنوعی تنها به حوزه معرفت‌شناسی محدود نمی‌ماند، بلکه به‌مثابه تلاشی برای بازتعریف نسبت میان دانستن و زیستن مطرح می‌شود؛ آن‌هم در شرایطی که فرایند شناخت از اختیار مستقیم انسان عبور کرده و در بسترهای فناورانه گسترش یافته است. نوآوری نظری این پژوهش در ارائه مدلی است که دانایی را از سطح حالت ذهنی فراتر برده و آن را به‌مثابه کنش اخلاقی- تفسیری در شبکه‌ای میان‌عاملی تبیین می‌کند. در این افق، حقیقت به‌عنوان رخدادی فهم می‌شود که در توان پاسخ‌گویی و شفافیت فرایند شناخت تحقق می‌یابد و دیگر صرفاً به‌مثابه واقعیتی بیرونی تعریف نمی‌شود.

مدل معرفت- اخلاق نشان می‌دهد که مشروعیت دانایی در جهان هوشمند بر پیوند سه مؤلفه استوار است؛ تبیین‌پذیری معرفتی، پاسخ‌گویی اخلاقی و تعامل تفسیری. بر پایه این قیاس، اگر دانایی از مسیر محاسبه و داده شکل می‌گیرد و اگر اخلاق بر توان پاسخ‌دادن استوار است، پس مسئولیت باید به ساختاری شبکه‌ای و میان‌عاملی انتقال یابد. این گزاره، هسته نوآورانه مدل را تشکیل می‌دهد و تبیین می‌کند که دانایی در عصر هوش مصنوعی زمانی معتبر است که بتواند پیامدهای اخلاقی خود را در سطح طراحی و تصمیم‌آشکار سازد. از منظر تحلیلی، مسئله اصلی هوش مصنوعی در فقدان ظرفیت شناختی خلاصه نمی‌شود، بلکه در نبود بُعد تفسیری و معنابخش نهفته است؛ بُعدی که بتواند داده را به تجربه اخلاقی پیوند دهد. ماشین در سطح استنتاج آماری عمل می‌کند، اما معنا را تجربه نمی‌کند. بدین‌سان، نقش انسان از جایگاه ناظر فراتر می‌رود و به‌عنوان میانجی معنا

بازتعریف می‌شود؛ عاملی که توانایی دارد داده را به فهم و فهم را به تعهد اخلاقی تبدیل کند. این همان لحظه‌ای است که معرفت از اطلاع جدا می‌شود و به مسئولیت بدل می‌گردد.

از دید هستی‌شناختی، دانایی از سطح ویژگی ذهن منفرد فراتر می‌رود و به‌مثابه رخدادی در تعامل میان انسان، سامانه و جامعه پدیدار می‌شود. از دید اخلاقی، صدق در نسبت با شفافیت و پاسخ‌گویی معنا می‌یابد و از دید تفسیری، فهم به میانجی گفت‌وگو میان انسان و ماشین تداوم می‌یابد. بدین‌گونه، فلسفه از مقام ناظر به جایگاه کنشگر بازمی‌گردد؛ کنشگری که در طراحی معرفت جهان مشارکت دارد. در نتیجه، سهم نظری این پژوهش در ارائه مدلی است که معرفت را در قالب «دانایی مسئول» بازمی‌فهمد؛ دانایی‌ای که در آن حقیقت، اخلاق و تفسیر به‌صورت هم‌زمان عمل می‌کنند. این مدل نشان می‌دهد که برای حفظ مشروعیت معرفتی در جهان هوشمند، باید اخلاق را از حاشیه فناوری به مرکز فرایند شناخت منتقل کرد. در این چشم‌انداز، آینده فلسفه در توانایی آن برای بازسازی مفاهیم حقیقت، معنا و فاعلیت در افق شبکه‌ای دانایی رقم می‌خورد؛ افقی که در آن دانستن به‌مثابه شکلی از زیستن اخلاقی فهم می‌شود.

References

- Abramoff, M. D., Lavin, P. T., Birch, M., Shah, N., & Folk, J. C. (2018). Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Medicine*, 1(1), 39.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2022). Machine bias. In *Ethics of data and analytics* (pp. 254–264). Auerbach Publications.
- Aristotle. (1994). *Posterior Analytics* (J. Barnes, Ed.). In *The Complete Works of Aristotle* (Vol. 1). Oxford: University of Oxford; Balliol College.
- Aristotle. (2000). *Nicomachean Ethics* (R. Crisp, Trans.). Cambridge: Cambridge University Press.
- Blanshard, B. (1939/2014). *The nature of thought: Volume I*. London: Routledge.
- Cartwright, N. (1983). *How the laws of physics lie*. OUP Oxford.
- Clark, A. (۲۰۰۸). *Supersizing the mind: Embodiment, action, and cognitive extension*. oxford university Press.
- Derrida, J. (1995). *The gift of death*. University of Chicago Press.
- Descartes, R. (1996). *Meditations on first philosophy* (J. Cottingham, Trans.). Cambridge: Cambridge University Press.
- Eubanks, V. (2025). *Automating inequality: How high-tech tools profile, police, and punish the poor*. Macmillan+ ORM.
- Floridi, L. (2011). *The philosophy of information*. Oxford: Oxford University Press.
- Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford: Oxford University Press.
- Foucault, M. (2013). *Archaeology of knowledge*. Routledge.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Gadamer, H. G. (2004). *Truth and method* (J. Weinsheimer & D. G. Marshall, Trans., Rev. ed.). London: Continuum.
- Goldman, A. I. (1986). *Epistemology and cognition*. harvard university Press.
- Goldman, A. I. (2010). *Why social epistemology is real epistemology*. Na.
- Habermas, J. (1998). Between facts and norms: An author's reflections. *Denv. UL Rev.*, 76, 937.
- Hedlund, M., & Persson, E. (2025). Distribution of responsibility for AI development: expert views. *AI & SOCIETY*, 1-13.

- Heidegger, M. (1962/2001). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Oxford: Blackwell.
- Heidegger, M. (1977). The question concerning technology. In *Technology and values: Essential readings* (pp. 99–113). New York: Harper.
- Husserl, E. (1983). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy: First book: General introduction to a pure phenomenology* (Vol. 2). Springer Science & Business Media.
- James, W. (1975). *Pragmatism and the Meaning of Truth*. CreateSpace Independent Publishing Platform.
- Johnson, D. G. (2015). Technology with no human responsibility? *Journal of Business Ethics*, 127(4), 707-715.
- Jonas, H. (1984). *The imperative of responsibility: In search of an ethics for the technological age*. University of Chicago press.
- Kant, I. (1998/1999). *Critique of pure reason* (P. Guyer & A. W. Wood, Eds.). Cambridge: Cambridge University Press.
- Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31-57.
- Longino, H. E. (2019). Science as social knowledge: Values and objectivity in scientific inquiry.
- Mill, J. S. (1998). *Utilitarianism* (R. Crisp, Ed.). Oxford: Oxford University Press.
- Moore, G. E. (1953/2014). *Some main problems of philosophy*. London: Routledge.
- O'neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. penguin UK.
- Plato. (2008). *Theaetetus* (S. Asscher & D. Widger, Eds.). Project Gutenberg. Retrieved from <https://www.gutenberg.org/ebooks/1726>
- Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the academy of marketing science*, 48(1), 137-141.
- Rawls, J. (1971). *A theory of justice*, Harvard. Press, Cambridge.
- Sartre, J. P. (1943/2003). *Being and nothingness* (H. Barnes, Trans.). London: Routledge.
- Sartre, J. P. (2007). *Existentialism is a humanism*. New Haven: Yale University Press.
- Stiegler, B. (1998). *Technics and time, 1: The fault of Epimetheus*. Stanford: Stanford University Press.
- Sunstein, C. R. (2018). *Republic: Divided democracy in the age of social media*.
- Wittgenstein, L. (2009). *Philosophical investigations*. John Wiley & Sons.