Quarterly Journal Science of Media Futures Studies
Vol. 1, Series. 1, No. 1, Spring 2026/ 109-137
DOI: https://doi.org/10.22034/mfs.2026.742358

The Future Landscape of Professional Ethics in the Age of Artificial Intelligence: A Framework for Requirements and Indicators of Professional Ethics in Media

Yousef Khojir *, Associate Professor of Communications, Department of Social Communications, Faculty of Culture and Communications, Soore International University, Tehran, Iran.*
Email: khojir@soore.ac.ir

Mahsa Khalat Abad *, PhD student in Communications, Faculty of Culture and Communications, Soore International University, Tehran, Iran. Email: m.khalatabad@soore.ac.ir*

Extended Abstract

Introduction: Given the rapid and undeniable penetration of artificial intelligence (AI) technologies into various levels of social structures-particularly within complex decision-making processes, big-data-driven analyses, and the fundamental transformation of media systems-addressing the behavioral, ethical, and functional implications of these systems has become inevitable. Although numerous studies in recent years have addressed general ethical principles and theoretical concepts in this field, a fundamental and evident gap remains in the scholarly literature: the absence of an operational, localized, and indicator-based framework capable of establishing a reliable bridge between abstract ethical principles and their implementation in professional practice.

This deficiency is particularly palpable within the specific context and operational requirements of our country's media ecosystem. Consequently, the central issue and primary motivation of this research is to conduct an in-depth analysis of the dimensions, identify the key components, and derive precise indicators of professional ethics in the realm of AI; by articulating the associated executive and operational requirements, the study aims to establish a robust foundation for fostering public trust and institutionalizing the responsible, ethical use of these emerging technologies within Iran's media landscape.

Methods: In terms of purpose, this study is applied, while regarding its nature and approach, it falls under the category of qualitative research utilizing thematic analysis. The study population consisted of academic elites and executive experts well-versed in professional ethics, media, and artificial intelligence. Sampling was conducted using a criterion-based purposive method combined with the snowball technique.

Data collection involved in-depth, semi-structured interviews with 17 specialists; initial signs of theoretical saturation emerged during the fourteenth interview, and the process continued until the seventeenth interview to ensure data adequacy. Data analysis was performed using MAXQDA software based on King's (1998) three-level coding model, wherein 166 extracted key phrases were transformed

into-and validated as-102 basic themes, 35 organizing themes, and ultimately 13 global themes.

Findings: Thematic analysis reveals that professional ethics in the realm of artificial intelligence is a highly multidimensional, structured, and inherently interdisciplinary concept. Far from being confined to purely technical and algorithmic variables, this field is deeply intertwined with complex legal, social, and human structures.

Indeed, the study's findings indicate that establishing a professional ethics framework requires a holistic approach that simultaneously addresses various operational and governance dimensions. A detailed examination categorizes the findings into five major themes. First is technological ethics and value-based algorithms, emphasizing the need to design transparent, supervisable, and human-centric systems to ensure alignment with human values.

Second is data governance and security, prioritizing privacy protection, the intelligent safeguarding of user data, and strict adherence to intellectual property rights. Third is algorithmic fairness and bias mitigation, focusing on the systematic refinement of data to prevent structural, social, cultural, and gender-based discrimination. Fourth is responsibility and accountability, mandating the transparent acceptance of legal and ethical obligations regarding automated decisions.

Finally, the fifth theme encompasses human capital development and media content ethics, addressing the enhancement of media literacy, the empowerment of ethical professionals, and the development of technological mechanisms to detect and curb fake news.

In the final stage of the analysis-aimed at prioritizing the indicators-it was determined that the components of security and privacy, social justice, and process transparency registered the highest frequency in the software-based analysis due to their foundational nature. Furthermore, an examination of the consensus among participating experts reveals that these three areas are regarded as the most urgent and critical requirements for AI ethics policymaking, while other dimensions fall into subsequent tiers of priority.

Conclusion: The findings of this research clearly emphasize that professional ethics in the age of artificial intelligence are not a peripheral consideration, but a fundamental and infrastructural prerequisite for achieving sustainable development and preserving and enhancing social capital within media systems. In fact, public trust in technology only emerges within the framework of ethical governance; failing to recognize this inseparable connection could transform the transformative capabilities of AI into tools for eroding society's symbolic capital.

The distinct achievement of this study, in contrast to dominant global literature, is its departure from purely abstract approaches toward the "localization of indicators." While international models primarily focus on universal principles, this research, by deeply understanding the specific requirements of Iran's cultural and media ecosystem, proposes an operational framework that acts as a "bridge," re-

pairing the existing gap between ethical theories and organizational behaviors. This model elevates the implementation of ethics from the level of theoretical statements to the level of actionable protocols and institutionalized organizational behaviors.

The present study warns that neglecting these ethical requirements is not merely a management gap; it can pave the way for profound social crises-crises stemming from widespread distrust in intelligent outputs, the targeted dissemination of distorted news, and algorithmic biases, which ultimately lead to the exacerbation of social divisions and the polarization of society.

Keywords: Ethics, professional ethics, digital ethics, artificial intelligence, Iran.

فصلنامه دانش آینده‌پژوهی رسانه
سال ۱ / شماره ۱/ شماره مسلسل ۱ / بهار ۱۴۰۵/ ۱۰۹ـ۱۳۷
Quarterly Journal Science of Media Futures Studies
Vol. 1, Series. 1, No. 1, Spring 2026/ 109-137
DOI: https://doi.org/10.22034/mfs.2026.742358

سیمای آینده اخلاق حرفه‌ای در عصر هوش مصنوعی؛ چارچوبی برای الزامات و شاخص‌های اخلاق حرفه‌ای در رسانه

یوسف خجیر[۱]، مهسا خلعت‌آباد[۲]

چکیده

این پژوهش با هدف واکاوی ابعاد، مؤلفه‌ها و شاخص‌های اخلاق حرفه‌ای در حوزه هوش مصنوعی، با تأکید بر تبیین الزامات اخلاقی این فناوری در سطح حرفه‌ای انجام شده است. با گسترش کاربرد هوش مصنوعی در فرایندهای تصمیم‌گیری، تحلیل داده و ارائه خدمات، توجه به اصول اخلاق حرفه‌ای به یکی از ضرورت‌های بنیادین در توسعه و بهره‌برداری از این فناوری تبدیل شده است. پژوهش حاضر از نظر ماهیت، کیفی و با استفاده از روش تحلیل مضمون انجام شده است. داده‌ها از طریق مصاحبه‌های عمیق نیمه‌ساختاریافته با ۱۷ نفر از خبرگان و نخبگان آشنا بااخلاق حرفه‌ای و هوش مصنوعی گردآوری شد و تحلیل داده‌ها براساس الگوی کدگذاری سه‌گانه کینگ (۱۹۹۸) صورت گرفت. یافته‌ها نشان داد که اخلاق حرفه‌ای در هوش مصنوعی مفهومی چندبعدی و میان‌رشته‌ای است که صرفاً به جنبه‌های فنی محدود نمی‌شود و با ابعاد حقوقی، اجتماعی و فرهنگی نیز پیوند دارد. از مهم‌ترین مؤلفه‌های شناسایی‌شده می‌توان به طراحی الگوریتم‌های اخلاق‌محور، حفظ حریم خصوصی و حفاظت از داده‌ها، رعایت عدالت و جلوگیری از تبعیض الگوریتمی، شفافیت و پاسخگویی در فرایندها و تصمیم‌های مبتنی بر الگوریتم و نیز تولید محتوای دقیق و معتبر اشاره کرد. نتایج نشان می‌دهد که این مؤلفه‌ها زیرساخت شکل‌گیری اعتماد عمومی نسبت به سامانه‌های هوشمند را فراهم می‌کنند و بی‌توجهی به آنها می‌تواند به گسترش بی‌اعتمادی، انتشار اطلاعات نادرست و تشدید نابرابری‌ها منجر شود. درمجموع، تأکید پژوهش بر این است که رعایت اصولی همچون شفافیت، عدالت، مسئولیت‌پذیری، حفاظت از حریم خصوصی و امنیت اطلاعات، همراه با تدوین سیاست‌ها و مقررات مناسب و آموزش نیروی انسانی متخصص، پیش‌شرط توسعه حرفه‌ای و پایدار هوش مصنوعی و استفاده مسئولانه از آن در جامعه به شمار می‌آید.

واژگان کلیدی

اخلاق، اخلاق حرفه‌ای، اخلاق دیجیتال، هوش مصنوعی، سواد رسانه‌ای، عدالت الگوریتمی، ایران.

تاریخ دریافت: ۱۴۰۴/۱۲/۰۵ تاریخ پذیرش: ۱۴۰۴/۱۲/۲۳

۱. دانشیار ارتباطات، گروه ارتباطات اجتماعی، دانشکدهٔ فرهنگ و ارتباطات، دانشگاه بین‌المللی سوره، تهران. ایران.
khojir@soore.ac.ir

۲. دانشجوی دکتری ارتباطات، دانشکدهٔ فرهنگ و ارتباطات، دانشگاه بین‌المللی سوره، تهران. ایران (نویسنده مسئول).
m.khalatabad@soore.ac.ir

مقدمه

اخلاق همواره به‌عنوان شالوده بنیادین شکل‌گیری جوامع و قوام تمدن‌ها، نقشی کلیدی در تنظیم مناسبات ارتباطی، حفظ انسجام اجتماعی و تولید اعتماد عمومی ایفا کرده است. با گذار به عصر ارتباطات شبکه‌ای و تحقق ایده «دهکده جهانی» مک لوهان (۱۹۶۲)، مرزهای ارتباطات انسانی دگرگون شد و زیست در بوم‌سازگان دیجیتال، دل‌مشغولی‌های هنجاری جدیدی را پدید آورد. در این میان، ظهور فناوری‌های برهم‌زننده، به‌ویژه هوش مصنوعی، ضرورت بسط اخلاق سنتی و شکل‌گیری «اخلاق دیجیتال» را به یک الزام گریزناپذیر تبدیل کرده است. هوش مصنوعی با ظرفیت‌های بی‌بدیل خود در تحلیل کلان‌داده‌ها و تولید خودکار محتوا، الگوهای روایت‌سازی، مدیریت تصویر و تعامل با افکار عمومی را به‌طور بنیادین بازتعریف کرده است؛ اما این گستردگی الگوریتمی، هم‌زمان چالش‌ها و پیامدهای پیش‌بینی‌نشده‌ای را در حوزه ارتباطات جمعی به همراه داشته است.

رسانه‌ها و پلتفرم‌های ارتباطی در تقاطع با هوش مصنوعی، با بحران‌های نوپدیدی نظیر سوگیری الگوریتمی، تولید هوشمند اطلاعات نادرست، جعل عمیق[1] و نقض پنهان حریم خصوصی کاربران مواجه‌اند. این پدیده‌ها نه‌تنها گفتمان عمومی و اعتماد نهادی را مستقیماً تهدید می‌کنند، بلکه نشان می‌دهند ابزارهای قانونی سلبی، به دلیل سرعت خیره‌کننده تغییرات فنّاورانه، برای برقراری نظم اخلاقی و صیانت از کرامت انسانی کفایت نمی‌کنند. ازاین‌رو، ضرورت استقرار نظام‌های خودتنظیم‌گر[2] و بسط «اخلاق حرفه‌ای» در این زیست‌بوم برجسته می‌شود. اخلاق حرفه‌ای رسانه در عصر هوش مصنوعی، مستلزم پایبندی داوطلبانه توسعه‌دهندگان فناوری و اصحاب رسانه به مجموعه‌ای از قواعد هنجاری است که رسالت اجتماعی و حقیقت‌جویی را فراتر از الزامات حداقلیِ قانون تضمین نماید با وجود اهمیت روزافزون این ضرورت هنجاری، بررسی ادبیات پژوهش حاکی از آن است که عمده دستورالعمل‌ها و چارچوب‌های اخلاقی توسعه‌یافته در زمینهٔ هوش مصنوعی، دارای ماهیتی جهان‌شمول و وارداتی هستند. در نظام ارتباطی فعلی، خلأ یک شیوه‌نامه اخلاقیِ مدون، بومی و مبتنی بر اجماع نخبگانی به‌شدت احساس می‌شود؛ چارچوبی که بتواند ضمن پاسخگویی به چالش‌های استقلال دیجیتال و صیانت از افکار عمومی در برابر تحریف روایی، باارزش‌های

1. Deepfakes
2. Self-regulation

فرهنگی و زمینه‌های اجتماعی همسو باشد؛ بنابراین، سؤال اصلی این پژوهش این است که معیارها، مؤلفه‌ها و سازوکارهای تحقق اخلاق حرفه‌ای در توسعه و کاربست هوش مصنوعی کدام‌اند؟ در پاسخ به این سؤال، مجموعه‌ای از سؤالات فرعی نیز مطرح شده‌اند که به تبیین دقیق‌تر ابعاد پژوهش کمک می‌کنند. ازجمله این سؤالات می‌توان به شناسایی اصول و ابعاد کلیدی اخلاق حرفه‌ای (نظیر شفافیت، عدالت و مسئولیت‌پذیری) در چرخه حیات سیستم‌های هوشمند، بررسی تأثیر رعایت این اصول بر کاهش سوگیری‌های الگوریتمی، حفظ حریم خصوصی و ارتقای اعتماد کاربران اشاره کرد. علاوه بر این، پژوهش به بررسی چالش‌ها و تنگناهای اخلاقی در طراحی و پیاده‌سازی الگوریتم‌ها و تعیین پیش‌نیازهای سیاستی، حقوقی، نهادی و آموزشی موردنیاز برای نهادینه‌سازی رفتار حرفه‌ای و مسئولانه در اکوسیستم هوش مصنوعی می‌پردازد.

پیشینه پژوهش

رشدانی (۱۴۰۴) در پژوهشی به بررسی پیش‌نیازهای چندوجهی پذیرش فناوری هوش مصنوعی در دستگاه‌های دولتی پرداخته و نتیجه می‌گیرد که پذیرش هوش مصنوعی صرفاً یک فرایند فناورانه نیست، بلکه به‌شدت تحت تأثیر عوامل فرهنگی، اخلاقی و ساختاری قرار دارد و در بعد اخلاقی و اجتماعی، مسائلی چون شفافیت تصمیمات الگوریتمی، حفظ حریم خصوصی، عدالت اطلاعاتی و مسئولیت‌پذیری در استفاده از داده‌ها از موانع یا محرک‌های اصلی پذیرش هوش مصنوعی در نهادهای عمومی به‌شمار می‌روند. قلعه‌نوئی (۱۴۰۴) در پژوهشی به بررسی جایگاه حریم خصوصی در چارچوب حقوق بشر پرداخت و نتیجه گرفت که نظام حقوق بشر برای مقابله مؤثر با چالش‌های نوظهور ناشی از فناوری‌های نوین نیازمند بازاندیشی و ایجاد چارچوب‌های الزام‌آور جدید است. درحالی‌که پژوهش حاضر با رویکردی آینده‌نگر به بررسی حریم خصوصی در جامعه و در بستر توسعه هوش مصنوعی در ایران می‌پردازد. خجیر و خلعت آباد (۱۴۰۳) در پژوهشی با عنوان «تدوین معیارهای اخلاق حرفه‌ای استفاده هوش مصنوعی در رسانه‌های ایران» به بررسی تعیین معیارهای اخلاق حرفه‌ای استفاده هوش مصنوعی در رسانه‌های ایران پرداخته و نتیجه می‌گیرند که مهم‌ترین معیار اخلاقی در استفاده هوش مصنوعی در رسانه‌های ایران شفافیت در الگوریتم‌ها و تصمیم‌گیری‌ها، عدالت و عدم تبعیض، حفاظت از حریم خصوصی، مسئولیت‌پذیری در برابر پیامدهای استفاده

از هوش مصنوعی و توسعه و پایداری است. رحمانی و همکاران (۱۴۰۳) در پژوهشی به این نتیجه دست یافتند که ترکیب هوش مصنوعی با اصول اخلاقی می‌تواند به بهبود فرایندهای تصمیم‌گیری استراتژیک و کاهش خطرات احتمالی منجر شود. فادیکانی و پارسیانا (۱۴۰۲) به این نتایج دست یافتند که این فناوری در ابعاد مختلف زندگی انسان نفوذ کرده، اما نگرانی‌هایی مانند حریم خصوصی، تبعیض در تصمیم‌گیری‌ها و عدم شفافیت الگوریتم‌ها را به وجود آورده است؛ که چالش‌ها نیازمند تدوین استانداردهای اخلاقی برای حفظ حقوق فردی و اجتماعی و کمک به توسعه پایدار و شکوفایی انسانی هستند.

الزوبی[1] و همکاران (۲۰۲۵) نیز به این نتایج دست یافتند که این رویکرد می‌تواند با ایجاد اختلال کنترل‌شده در داده‌ها از افشای اطلاعات حساس کاربران جلوگیری کند و درعین‌حال امکان استفاده از داده‌ها برای تحلیل‌های هوشمند را فراهم سازد؛ هرچند اجرای مؤثر آن همچنان با چالش‌های فنی و محدودیت‌های عملی همراه است. یافته‌های کوروبنکو[2] و همکاران (۲۰۲۵) نشان می‌دهد که توجه به مؤلفه‌هایی همچون امنیت اطلاعات، شفافیت و حفاظت از حریم خصوصی کاربران از الزامات اساسی در توسعه سامانه‌های هوش مصنوعی اخلاق‌محور محسوب می‌شود. بیلریس[3] و همکاران (۲۰۲۵) نیز در پژوهشی با عنوان «حریم خصوصی در عصر هوش مصنوعی: طبقه‌بندی خطرات داده‌ها»[4] به تحلیل نظام‌مند تهدیدهای مرتبط با داده‌ها در سامانه‌های هوش مصنوعی پرداخته‌اند. این پژوهش با ارائه یک چارچوب طبقه‌بندی نشان می‌دهد که ترکیب و تحلیل گسترده داده‌ها می‌تواند خطر افشای اطلاعات شخصی را افزایش دهد و برای مدیریت این خطرات باید سیاست‌های چندلایه در حوزه مدیریت و حفاظت از داده‌ها طراحی شود. مطالعات آکاکپو[5] و همکاران (۲۰۲۴) نشان می‌دهد که پیشرفت‌های سریع در تحلیل داده‌ها چالش‌های جدیدی برای حفاظت از اطلاعات شخصی ایجاد کرده و ضرورت ایجاد تعادل میان نوآوری فناورانه و صیانت از حقوق کاربران را برجسته ساخته است. مارتین و زیمرمن[6] (۲۰۲۴) در پژوهشی بر ضرورت تدوین چارچوب‌های حقوقی و اخلاقی جدید برای مدیریت داده‌ها و کاهش خطر

1. Alzoubi
2. Korobenko
3. Billiris
4. Privacy in the age of AI: A taxonomy of data risks
5. Akakpo
6. Martin & Zimmermann

نقض حریم خصوصی در عصر هوش مصنوعی تأکید می‌کند. کلوژ کوریا[1] و همکاران (۲۰۲۴) به این نتایج دست یافتند که اصولی مانند شفافیت، عدالت، مسئولیت‌پذیری و احترام به حریم خصوصی از مهم‌ترین مبانی اخلاقی در توسعه و کاربرد فناوری‌های هوش مصنوعی به شمار می‌روند. همچنین فلوریدی[2] (۲۰۲۳) در کتاب خود با عنوان «اخلاق هوش مصنوعی: اصول، چالش‌ها و فرصت‌ها»[3] به بررسی دو هدف اصلی اخلاق هوش مصنوعی ازجمله: تحلیل نظری که شامل بررسی گذشته، حال و آینده هوش مصنوعی و تأکید بر جدایی آژانس و اطلاعات است و بررسی پیامدها و کاربردها که به چالش‌های اخلاقی، خطرات بالقوه و شیوه‌های خوب استفاده از این فناوری پرداخته است. کتاب همچنین تأثیرات هوش مصنوعی بر محیط‌زیست و نقش آن در مبارزه با تغییرات آب و هوایی را بررسی کرده و درنهایت بر اهمیت هم‌افزایی بین محیط‌زیست و فناوری‌های دیجیتال برای ایجاد جامعه‌ای بهتر و زیست‌کره‌ای سالم‌تر تأکید می‌کند.

با بررسی و مرور نظام‌مند پیشینه می‌توان دریافت که مطالعات صورت گرفته در حوزه اخلاق هوش مصنوعی، تاکنون عمدتاً بر ابعاد حقوقی، اجتماعی، مدیریتی و فناورانه این پدیده متمرکز بوده‌اند. در سطح داخلی، بخشی از پژوهش‌ها بر نقش اخلاق در حکمرانی، مدیریت و رسانه‌ها تأکید ورزیده‌اند؛ در سطح بین‌المللی نیز آثار متعددی به واکاوی اهمیت حفاظت از حریم خصوصی، امنیت داده‌ها، طراحی چارچوب‌های ایمن برای توسعه سامانه‌ها و ضرورت تدوین سیاست‌ها و مقررات جدید برای صیانت از اطلاعات شخصی پرداخته‌اند. درواقع، تمرکز اصلی دانش موجود بر زیرساخت‌ها، قوانین و سیاست‌گذاری‌های کلان بوده است. باوجود غنای مطالعات یادشده در سطوح کلان و فناورانه، یک خلأ پژوهشی جدی و عمیق در ادبیات موجود به چشم می‌خورد: غفلت از بُعد اخلاق حرفه‌ای و رفتار فردی متخصصان. به بیان دقیق‌تر، مطالعات پیشین عمدتاً از منظر سیاست‌گذار یا طراح سیستم به موضوع نگریسته‌اند و کمتر به این پرسش پرداخته‌اند که اخلاق در لایه کنش حرفه‌ای و در محیط کار شاغلان هوش مصنوعی چگونه باید نمود یابد. در ادبیات فعلی، الگوی مدونی که الزامات اخلاقی را به شاخص‌های رفتاری و استانداردهای شغلی برای متخصصان این حوزه تبدیل کند، وجود ندارد. بر این اساس، پژوهش حاضر باهدف

1. Kluge Corrêa
2. Floridi
3. The ethics of artificial intelligence: Principles, challenges, and opportunities

پوشش مستقیم این خلأ معرفتی تدوین شده است. این مطالعه در پی آن است که با عبور از مباحث کلی حکمرانی و حقوقی، مستقیماً بر استفاده اخلاقی هوش مصنوعی در حرفه و اخلاق حرفه‌ای متمرکز شود. هدف بنیادین این پژوهش، استخراج و تدوین الزامات و شاخص‌های اخلاق حرفه‌ای است؛ به‌گونه‌ای که فاصله میان سیاست‌های انتزاعی کلان و عمل روزمره متخصصان این حوزه کاهش یابد و نقشه راهی کاربردی برای رفتارهای اخلاقی در اکوسیستم هوش مصنوعی ارائه گردد.

ادبیات پژوهش و چارچوب مفهومی

ادبیات پژوهش

چیستی اخلاق

اخلاق به مثابه یک نظام هنجاری و درونی، مجموعه‌ای از اصول و ارزش‌هاست که تنظیم‌گر رفتارهای انسانی بوده و به ارتقای کیفیت زندگی و رفاه فردی و اجتماعی می‌انجامد. این مفهوم که فرایندهایی عینی و معنوی را در برمی‌گیرد، نقش بسزایی در رشد و تعادل ویژگی‌های انسانی ایفا کرده و قادر است بروز رفتارهای مطلوب را بدون نیاز به محرک‌ها و الزام‌های بیرونی تضمین نماید (حسینی و عباسی، ۱۳۹۱؛ صلواتی و همکاران، ۱۳۹۲). در ساحت آکادمیک، علم اخلاق به‌عنوان شاخه‌ای بنیادین از علوم انسانی که به تعیین و ارزش‌گذاری رفتارهای انسانی می‌پردازد، در سه قلمروِ «فلسفه اخلاق» (بررسی پرسش‌های بنیادین)، «اخلاق نظری» (تحلیل صفات و انتظارات اخلاقی) و «اخلاق عملی» (تربیت و نهادینه‌سازی رفتارهای اخلاقی در جامعه) دسته‌بندی می‌شود (کرمپوریان و همکاران، ۱۳۹۲؛ صلواتی و همکاران، ۱۳۹۲). درنهایت، اخلاق پیوندی ناگسستنی باارزش‌ها دارد و به‌عنوان ابزاری کارآمد، زمینه‌ساز تجلی عملیِ این ارزش‌هاست؛ قواعدی که متناسب با بسترهای فرهنگی و دوره‌های زمانی مختلف، تنوع و پویایی خاص خود را دارند (حسینی و عباسی، ۱۳۹۱).

چیستی اخلاق حرفه‌ای

اخلاق حرفه‌ای تبلور و امتداد اصول اخلاق عمومی در بستر مشاغل و محیط‌های کاری است که به تبیین و حل‌وفصل مسائل و چالش‌های هنجاریِ مختص به هر حرفه می‌پردازد. این مفهوم دربرگیرنده مجموعه‌ای از قواعد، اصول و استانداردهای رفتاری است که افراد نه از سر اجبار قانونی، بلکه به صورت داوطلبانه و مبتنی بر تعهد درونی خویش را ملزم به رعایت آنها می‌دانند. رسالت اصلی اخلاق حرفه‌ای،

مدیریت رفتارها و تبیین مسئولیت‌های اخلاقی کارکنان و سازمان‌ها در قبال جامعه و ذی‌نفعان است. پایبندی به این اصول اخلاقی در محیط کار، به‌عنوان یک محرک بنیادین عمل کرده و ضمن تضمین سلامت محیط کاری، به بهبود ارتباطات انسانی، افزایش اعتماد و درنهایت ارتقای بهره‌وری و کارآمدی سازمانی منجر می‌شود (امین‌بیدختی و مردانی، ۱۳۹۴).

چیستی هوش مصنوعی

هوش مصنوعی به مجموعه‌ای از فناوری‌ها و الگوریتم‌ها اطلاق می‌شود که به ماشین‌ها این امکان را می‌دهد تا وظایف مشابه انسان‌ها را انجام دهند (راسل و نورویگ، ۲۰۲۰). این فناوری‌ها شامل یادگیری ماشین، یادگیری عمیق و پردازش زبان طبیعی هستند که به سیستم‌ها اجازه می‌دهند از داده‌ها یاد بگیرند و تصمیم‌گیری کنند. در دنیای امروز، هوش مصنوعی در حوزه‌های مختلفی ازجمله پزشکی، خودروسازی و رسانه‌ها مورداستفاده قرار می‌گیرد و به‌ویژه در تولید و تحلیل محتوا نقش مهمی ایفا می‌کند. درمجموع، هوش مصنوعی نوعی فناوری یا قطعی رایانه‌ای از وب معنایی (وب ۳) است که برگرفته و تقلیدی از هوش انسانی است و قابلیت ایجاد، تغییر و بهبود در فرایندها و قدرت تفکر و تصمیم‌گیری، برنامه‌ریزی و اجرا را دارد (خجیر، ۱۴۰۳: ۲۳).

چیستی اخلاق حرفه‌ای هوش مصنوعی

اخلاق هوش مصنوعی حوزه‌ای مطالعاتی و هنجاری است که به تعیین و تبیین اصول اخلاقی در تمامی مراحل طراحی، توسعه و کاربست سیستم‌های هوشمند پرداخته و هدف غایی آن، اولویت‌بخشی به رفاه انسانی و اجتماعی است. این حوزه تلاش می‌کند با ایجاد چارچوب‌ها و شیوه‌نامه‌های مدون (کرمپوریان و همکاران، ۱۳۹۲)، به مقابله با چالش‌ها و آسیب‌های نوپدید این فناوری نظیر نقض حریم خصوصی، انتشار اخبار جعلی، سوگیری‌های الگوریتمی و تبعیض‌های نژادی و جنسیتی ناشی از تحلیل داده‌های و فروش اطلاعات کاربران بپردازد. درواقع، اخلاق هوش مصنوعی پاسخی است به نگرانی‌های موجود در محیط‌های دیجیتال و رسانه‌ای که در آن، ضمن پذیرش ظرفیت‌های مثبت فناوری در ارتقای کیفیت محتوا و تسهیل تعاملات، بر ضرورت تربیت نیروی انسانی متخصص و پایبند به اصول اخلاقی تأکید می‌ورزد تا از پیامدهای مخربی همچون انحصار اطلاعاتی و کاهش تحمل‌پذیری در برابر نظرات مخالف پیشگیری شده و فناوری در مسیر حل معضلات اجتماعی هدایت شود.

چارچوب مفهومی

اخـلاق حرفـه‌ای در هوش مصنوعـی به‌عنـوان یـک نظـام هنجـاری و خودتنظیم‌گـر، بـه تبییـن تعهـدات، مسـئولیت‌ها و کدهـای رفتـاریِ متخصصـان، طراحـان و کارگـزارانِ درگیـر در چرخـه حیـات سیسـتم‌های هوشـمند (از ایـده تـا کاربسـت) می‌پـردازد. بـا توجـه بـه نفـوذ دگرگون‌کننـده ایـن فنـاوری در لایه‌هـای مختلـف جامعـه، رسـانه و اقتصـاد، صورت‌بنـدی یـک چارچـوب مفهومـی جامـع کـه نـه صرفـاً بـر خـود فنـاوری، بلکـه بـر «عاملیـت انسـانی و سـازمانی» در محیط‌هـای کاری تمرکـز داشـته باشـد، ضرورتـی انکارناپذیـر اسـت. ایـن چارچـوب بـر پایـه مؤلفه‌هایـی چندوجهـی بنـا شـده اسـت کـه اسـتانداردهای رفتـار حرفـه‌ای را در تقاطـع فنـاوری، حقـوق و علـوم ارتباطـات تعریـف می‌کننـد.

در سـطح عملکـرد فنـی و طراحـی الگوریتمـی، مؤلفه‌هـای «شـفافیت و توضیح‌پذیـری» به‌عنـوان یـک رسـالت حرفـه‌ای در کانـون توجـه قـرار دارنـد. توسـعه‌دهندگان و کارگـزاران ایـن حـوزه تعهـد اخلاقـی و شـغلی دارنـد کـه منطـقِ «جعبه‌سـیاه» را بـه چالـش کشـیده و سیسـتم‌هایی توسـعه دهنـد کـه چرایـی اتخـاذ تصمیمـات آنهـا بـرای ذی‌نفعـان قابـل درک باشـد (Miller, 2019)؛ ایـن مهـم به‌ویـژه بـرای متخصصانـی کـه در حوزه‌هـای حساسـی چـون روزنامه‌نـگاری داده، پزشـکی و حقـوق فعالیـت می‌کننـد، یـک اسـتاندارد قطعـی اسـت (Lipton, 2018). همـگام بـا ایـن مؤلفـه، پایبنـدی حرفـه‌ای بـه «عدالـت و عـدم تبعیـض» قـرار دارد. پایـش مسـتمر داده‌هـا بـرای شناسـایی و رفـع سـوگیری‌های پنهـان اجتماعـی یـا نـژادی، وظیفـه ذاتـی متخصصـان اسـت تـا از بازتولیـد نابرابری‌هـا جلوگیـری کـرده و نتایجـی منصفانـه را تضمیـن نماینـد (Obermeyer, 2019). عـلاوه بـر ایـن، تأمیـن «ایمنـی و امنیـت» سیسـتم‌ها و صیانـت از کاربـران در برابـر خطاهـای پردازشـی، بخـش جدایی‌ناپذیـری از تعهـدات کیفـی در اخـلاق حرفـه‌ای محسـوب می‌شـود (Brundage et al., 2018).

در سـطح حکمرانـی سـازمانی و تعامـل بـا کاربر، «حریـم خصوصـی» و «مسـئولیت‌پذیری» نقـش محـوری ایفـا می‌کننـد. رعایـت اصـول اخـلاق حرفـه‌ای ایجـاب می‌کنـد کـه سـازمان‌ها و شـاغلین ایـن حـوزه، فراتـر از الزامـات قانونـی، احتـرام بـه حقـوق داده‌ای افـراد را در دسـتورالعمل‌های درون‌سـازمانی خـود نهادینـه کننـد و شـفافیت لازم را در خصـوص نحـوه پـردازش اطلاعـات شـخصی ارائـه دهنـد (Zuboff, 2019). همزمـان، پذیـرش مسـئولیت‌پذیری سـازمانی بـرای تعییـن پاسـخگو در صـورت بـروز خطـا و آسـیب ناشـی از سیسـتم‌های هوش مصنوعـی (Gogoll & Müller, 2017). همچنیـن تدویـن کدهـای اخلاقیِ درون‌حرفـه‌ای بـرای نظـارت بـر عملکـرد کارکنـان، از ارکان اصلـی حاکمیـت شـرکتی در ایـن

عصـر بـه شـمار می‌آیـد (Jobin, Ienca, & Andorno, 2019).

در سـطح کلان و در سـاحت مسـئولیت اجتماعیِ حرفه‌هـا، بسـط هوش مصنوعـی نیازمنـد آینده‌نگـری اخلاقـی اسـت. توجـه بـه «تأثیـرات اجتماعـی و اقتصـادی» فنـاوری، نظیـر دگرگونـی در بـازار کار و لـزوم ارتقـای مهارت‌هـای نیـروی انسـانی، رسـالتی بـر دوش مدیـران و سیاسـت‌گذاران سـازمانی اسـت تـا از تشـدید نابرابری‌هـا جلوگیـری کننـد (Bessen, 2019). همچنیـن، امتنـاع حرفـه‌ای از مشـارکت در توسـعه «کاربردهـای مخـرب و ناقـض حریـم خصوصـی»، همچـون ابزارهـای نظـارت جمعـی غیراخلاقـی، آزمونـی بـرای پایبنـدی بـه حقـوق بشـر در محیـط کار اسـت (Zuboff, 2019). درنهایـت، ایـن چارچـوب بـا رویکـرد «توسـعه پایـدار»، متخصصـان را ملـزم می‌سـازد تـا در طراحـی زیرسـاخت‌های پردازشـی، پیامدهـای زیسـت‌محیطی را ارزیابـی و مدیریت نماینـد (Gonzalez et al., 2020).

درمجمـوع، اخـلاق حرفـه‌ای در هوش‌مصنوعـی پدیـده‌ای تقاطعـی اسـت کـه نهادینه‌سـازی آن، مسـتلزم گـذر از قوانیـنِ صرفـاً دسـتوری بـه سـمت کدهـای رفتـاری و نظام‌هـای خودتنظیم‌گـر شـغلی اسـت. مدیریـت ایـن حـوزه نیازمنـد هم‌افزایـی میان‌رشـته‌ای میـان متخصصـان فنـاوری، علـوم ارتباطـات و سیاسـت‌گذاری اسـت تـا ضمـن بهره‌منـدی از ظرفیت‌هـای هوش‌مصنوعـی، پایبنـدی کارگـزاران بـه منافـع عمومـی و کرامـت انسـانی تضمیـن گـردد.

روش پژوهش

ایـن پژوهـش ازنظـر ماهیت، کیفی و روش آن تحلیل مضمون اسـت. بـرای گردآوری اطلاعات از مصاحبه‌هـای نیمه‌سـاختارمند عمیـق بهـره گرفتـه شـد و جهـت پـردازش داده‌هـا و حفـظ جزئیـات و کلیـت مفاهیـم، از الگـوی کدگـذاری سـه‌گانه کینـگ (۱۹۹۸) اسـتفاده گردیـد.

جامعـه آمـاری ایـن مطالعـه شـامل خبـرگان و نخبـگان دارای اشـراف کامـل بـه حـوزه اخـلاق حرفـه‌ای رسـانه و هوش مصنوعـی اسـت. به‌منظـور گزینـش مشـارکت‌کنندگان، از نمونه‌گیـری هدفمنـد مبتنـی بـر معیـار[1] بـا بهره‌گیـری تکمیلـی از روش گلوله‌برفـی اسـتفاده شـد. ملاک‌هـای ورود بـه پژوهـش به‌تفکیـک بدیـن شـرح تعییـن گردیـد: نخبـگان دانشـگاهی می‌بایسـت دارای مـدرک دکتـری و حداقـل ۵ سـال سـابقه تدریـس تخصصـی در ایـن حـوزه باشـند؛ همچنیـن خبـرگان اجرایـی ملـزم بـه داشـتن مـدرک کارشناسی‌ارشـد و بالاتـر و حداقل ۳ سـال سـابقه فعالیـت مسـتقیم در حـوزه فناوری‌هـای نویـن رسـانه‌ای بودنـد. بـر این اسـاس،

1. Criterion-based Purposive Sampling

درمجموع ۷ استاد دانشگاه و ۱۰ متخصص حوزه اخلاق رسانه‌ای و هوش مصنوعی به‌عنوان مشارکت‌کننده انتخاب شدند. فرایند گردآوری داده‌ها تا حصول اشباع نظری تداوم یافت؛ به‌طوری‌که نشانه‌های اشباع و عدم ظهور کدهای جدید در مصاحبه چهاردهم پدیدار شد، اما جهت حصول اطمینان از عمق‌بخشی به یافته‌ها، مصاحبه‌ها تا مشارکت‌کننده هفدهم ادامه یافت. تمامی مصاحبه‌ها به صورت حضوری و مجازی با میانگین زمان ۶۰ دقیقه انجام، ضبط و پس از پیاده‌سازی، وارد چرخه کدگذاری و تحلیل نهایی شدند. افراد مشارکت‌کننده در فرایند مصاحبه در «جدول ۱» آمده است.

رعایت قابلیت اعتماد و اعتبار این پژوهش براساس رویکرد کرسول (۲۰۰۳) صورت گرفت. بدین‌صورت که ابتدا به فراهم کردن اطلاعات کافی و در دسترس پرداخته شد؛ سپس به اعتبار سنجی داده‌های جمع‌آوری‌شده و بعد نتایج را تکرار و تأیید آنان پرداخته و سپس به تأیید تفسیر داده‌های خود و گرفتن تأیید از افراد دیگر، داده‌ها را منطبق و تفسیر نمودیم؛ تمام مطالب به‌دست‌آمده را تفسیر و تحلیل کردیم.

جدول ۱. معرفی مشارکت‌کنندگان در مصاحبه

تحصیلات	مرتبه علمی/ تخصص	مدت آشنایی با موضوع
دکتری علوم ارتباطات	استاد ارتباطات/ رئیس هیئت‌مدیره انجمن سواد رسانه‌ای	۱۰ سال
دکتری هوش مصنوعی	استادیار/ عضو کمیته علمی ـ کنفرانس بین‌المللی وب‌پژوهی	۱۰ سال
دکتری مدیریت فناوری اطلاعات	استاد تمام فناوری اطلاعات/ متخصص هوش مصنوعی و استاد دانشگاه	۱۰ سال
دکتری فلسفه فناوری	استادیار فلسفه/ عضو هیئت‌علمی دانشگاه آزاد و سرپرست دانشکدهٔ مدیریت	۵ سال
دکتری علوم ارتباطات	استادیار ارتباطات/ رئیس انجمن ایرانی مطالعات جامعه اطلاعاتی	۱۰ سال
متخصص اخلاق و تربیت اسلامی	سطح ۴ حوزه/ پژوهشگر در حوزه اخلاق اسلامی	۵ سال
دکتری فلسفه فیزیک و سطح ۴ فقه و اصول حوزه علمیه	استادیار/ فعال حوزه‌های فلسفه، اخلاق، فقه و حقوق فناوری	۱۰ سال
دکتری تخصصی بیوتکنولوژی پزشکی	دانشیار/ رئیس مرکز پژوهش‌های مجلس،	۵ سال

تحصیلات	مرتبه علمی/ تخصص	مدت آشنایی با موضوع
دکتری علوم کامپیوتر	استادیار/ فعال حوزه یادگیری ماشین و هوش مصنوعی	۱۰ سال
دکتری مدیریت دولتی	استاد تمام/ عضو هیئت تحریریه در مجله علوم مدیریت و آموزش‌وپرورش و مجله بین‌المللی علوم مدیریت و اقتصاد	۷ سال
دکتری علوم ارتباطات	استاد مدعو/ مدیر پژوهش فضای مجازی و رسانه‌های جدید مرکز تحقیقات	۵ سال
دکتری علوم ارتباطات	دکتری/ مدرس فنّاوری ارتباطات دانشکدهٔ فرهنگ و ارتباطات سوره	۵ سال
دکتری فلسفه و فناوری	دکتری/ مدرس و پژوهش‌گر دانشگاه‌ها و مراکز تحقیقاتی	۵ سال
دکتری حقوق فناوری	دکتری/ پژوهشگر و فعال حوزه فناوری و حقوق	۵ سال
دکتری علوم کامپیوتر	دکتری/ متخصص علوم کامپیوتر و فعال در زمینهٔ سیاست‌گذاری‌های فناوری	۵ سال
فوق‌لیسانس مهندسی فناوری اطلاعات	فعالیت تخصصی/ فعال حوزه فناوری اطلاعات خبرگزاری جمهوری اسلامی	۲۴ سال
فوق‌لیسانس مهندسی کامپیوتر	فعالیت تخصصی/ کارشناس فناوری و مشاور پروژه‌های هوش مصنوعی	۱۵ سال

یافته‌های پژوهش

فرایند تحلیل داده‌ها با بهره‌گیری از نرم‌افزار MAXQDA10 و با طی کردن مراحل کدگذاری سه‌گانه کینگ (۱۹۹۸) انجام شد. در گام نخست، با بررسی خط‌به‌خط متن مصاحبه‌ها، ۱۶۶ عبارت کلیدی (واحد تحلیل) استخراج گردید. سپس، این عبارات از طریق فرایند کدگذاری باز، بازبینی و با ادغام موارد هم‌پوشان و حذف کدهای تکراری، به ۱۰۲ مضمون پایه تقلیل یافتند. در گام بعدی، باهدف انتزاعی‌تر کردن داده‌ها و شناسایی الگوهای تکرارشونده، مضامین پایه با توجه به شباهت‌های مفهومی و ارتباط منطقی، دسته‌بندی شدند؛ این فرایند کدگذاری محوری منجر به تشکیل ۳۵ مضمون سازمان‌دهنده گردید. درنهایت، با تلفیق و بررسی مضامین سازمان‌دهنده در سطحی بالاتر از انتزاع، ۱۳ مضمون فراگیر به‌عنوان شبکه مضامین نهایی استخراج شد که منعکس‌کننده ساختار کلان اخلاق حرفه‌ای در هوش مصنوعی است (جدول ۲).

جدول ۲. مضامین ابعاد اخلاق حرفه‌ای در هوش مصنوعی

مضمون پایه	مضمون سازمان دهنده	مضمون فراگیر
توسعه هوش مصنوعی نیازمند طرح و تبادل مستمر چالش‌های اخلاقی میان پژوهشگران و سیاست‌گذاران است.	اشتراک‌گذاری مسائل اخلاقی	اخلاق فناورانه
استفاده مؤثر از هوش مصنوعی مستلزم تعامل و همکاری سازنده میان انسان و سامانه‌های هوشمند است.	همکاری انسان و هوش مصنوعی	
سامانه‌های هوش مصنوعی باید به‌گونه‌ای طراحی شوند که در کنار انسان به‌عنوان همکار هوشمند عمل کنند.	تقویت کار تیمی انسان و هوش مصنوعی	
حفاظت از داده‌های شخصی کاربران یکی از اصول اساسی در کاربرد مسئولانه هوش مصنوعی است.	حفاظت از داده‌های کاربران	بهبود امنیت سایبری
ایجاد سازوکارهای امنیتی برای جلوگیری از سوءاستفاده از اطلاعات کاربران ضروری است.	حریم خصوصی و امنیت داده	
هوش مصنوعی می‌تواند در شناسایی و مقابله با اخبار جعلی نقش مؤثری داشته باشد.	تشخیص اخبار جعلی	
کاربرد صلح‌آمیز هوش مصنوعی می‌تواند به کاهش تنش‌های امنیتی کمک کند.	کاهش نگرانی‌های امنیتی و نظامی در سطح جهانی	
استفاده از داده‌ها و الگوریتم‌ها باید در چارچوب قوانین مالکیت معنوی انجام شود.	قوانین و مقررات مالکیت معنوی	حکمرانی و مدیریت دولتی
حقوق تولیدکنندگان داده و محتوای علمی و فرهنگی باید در توسعه هوش مصنوعی حفظ شود.	مالکیت فکری	
شفافیت در عملکرد سامانه‌های هوش مصنوعی موجب افزایش اعتماد عمومی می‌شود.	گزارش‌دهی شفاف	
مدل‌های هوش مصنوعی باید تا حد امکان قابل تفسیر و قابل‌درک باشند.	تفسیرپذیری مدل‌های پیچیده	
مسئولیت پیامدهای تصمیم‌های هوش مصنوعی باید مشخص و قابل‌پیگیری باشد.	مسئولیت تصمیم‌های مبتنی بر AI	
استفاده از فناوری باید از افزایش شکاف‌های اجتماعی جلوگیری کند.	جلوگیری از شکاف و چنددستگی	
پالایش داده‌ها برای حذف داده‌های معیوب در عملکرد سیستم‌های هوش مصنوعی ضروری است.	حذف داده‌های معیوب	
مدیریت مناسب جریان اطلاعات می‌تواند از بروز آسیب‌های اجتماعی جلوگیری کند.	مدیریت جریان اطلاعات	

مضمون پایه	مضمون سازمان دهنده	مضمون فراگیر
توسعه هوش مصنوعی باید باارزش‌های فرهنگی و اجتماعی جامعه سازگار باشد.	توجه به ابعاد مختلف در توسعه	لزوم توسعه پایدار
نظارت انسانی برای کنترل عملکرد و پیامدهای تصمیم‌های هوش مصنوعی ضروری است.	توسعه پایدار و پاسخگویی	
توسعه الگوریتم‌ها باید هم‌زمان کارایی فنی و اصول اخلاقی را رعایت کند.	بهبود عملکرد در چارچوب ارزش‌های انسانی	
کاربرد هوش مصنوعی نباید موجب نقض کرامت انسانی شود.	حفظ کرامت و شرافت انسانی	طراحی اخلاق‌محور الگوریتم‌ها
الگوریتم‌های هوش مصنوعی باید به‌طور مستمر مورد آزمون و ارزیابی قرار گیرند.	آزمون و ارزیابی مداوم الگوریتم‌ها	
توسعه سیستم‌های هوش مصنوعی باید باهدف افزایش دقت و قابلیت اعتماد انجام شود.	بهبود عملکرد سیستم‌ها	
سامانه‌های هوش مصنوعی باید خدمات خود را بدون تبعیض ارائه دهند.	ارائه خدمات بدون تعصب	عدالت الگوریتمی
در طراحی فناوری باید تنوع فرهنگی و اجتماعی کاربران در نظر گرفته شود.	توجه به تنوع و شمول اجتماعی	
توسعه‌دهندگان باید سوگیری‌های احتمالی در الگوریتم‌ها را کاهش دهند.	کاهش سوگیری در الگوریتم‌ها	
توسعه پایدار هوش مصنوعی نیازمند آموزش نیروی انسانی متخصص است.	تربیت نیروی انسانی متخصص	توسعه سرمایه انسانی
کاربرد هوش مصنوعی باید به حفظ سلامت فضای رسانه‌ای کمک کند.	مدیریت هوشمند اطلاعات	نگاه نوآورانه به هوش مصنوعی
پالایش داده‌ها برای افزایش دقت عملکرد سیستم‌های هوش مصنوعی ضروری است.	افزایش دقت داده‌ها	توسعه فناوری هوشمند
سامانه‌های هوش مصنوعی باید به تولید محتوای صحیح و قابل‌اعتماد کمک کنند.	تولید محتوای دقیق و معتبر	
هوش مصنوعی می‌تواند با خودکارسازی برخی وظایف بهره‌وری را افزایش دهد	خودکار سازی وظایف	
آموزش فناوری باید از ایجاد فاصله میان نسل‌ها جلوگیری کند.	جلوگیری از شکاف نسلی	همگام‌سازی ارزش‌ها
گسترش هوش مصنوعی می‌تواند موجب تحول در ساختارهای فرهنگی و اجتماعی شود	تحول در ساختارهای فرهنگی و اجتماعی	

مضمون پایه	مضمون سازمان دهنده	مضمون فراگیر
شناسایی و مدیریت خطرات فناوری از الزامات توسعه مسئولانه آن است.	پیشگیری از خطرات و چالش‌های فناوری	اطلاع‌رسانی و مدیریت ریسک
اطلاع‌رسانی شفاف می‌تواند به رفع نگرانی‌های عمومی درباره هوش مصنوعی کمک کند.	پاسخگویی به ابهامات افکار عمومی	آگاهی‌بخشی عمومی
مشارکت شهروندان در استفاده و سیاست‌گذاری فناوری اهمیت دارد.	تشویق مشارکت اجتماعی	
توسعه هوش مصنوعی باید بر اساس نیازهای واقعی جامعه انجام شود.	توجه به نیازهای واقعی جامعه	ارتقای سواد رسانه‌ای

اخلاق در هوش مصنوعی به‌عنوان یکی از ابعاد اساسی توسعه فناوری‌های نوین، مستلزم توجه هم‌زمان به ملاحظات اخلاقی، حقوقی، اجتماعی و فناورانه است. در این چارچوب، اشتراک‌گذاری مسائل اخلاقی میان پژوهشگران، سیاست‌گذاران و توسعه‌دهندگان فناوری از اهمیت ویژه‌ای برخوردار است؛ زیرا طرح و تبادل مستمر چالش‌های اخلاقی می‌تواند به شکل‌گیری رویکردی مسئولانه در طراحی و کاربرد سامانه‌های هوشمند کمک کند. در کنار این موضوع، همکاری و کار تیمی میان انسان و سامانه‌های هوش مصنوعی نیز از عناصر مهم اخلاق فناورانه محسوب می‌شود؛ به‌گونه‌ای که استفاده مؤثر از این فناوری مستلزم تعامل سازنده میان انسان و فناوری و طراحی سیستم‌هایی است که در کنار انسان به‌عنوان همکار هوشمند فعالیت کنند (مشارکت‌کننده شماره ۲). از منظر امنیت و حریم خصوصی، حفاظت از داده‌های کاربران و تضمین امنیت اطلاعات از اصول بنیادین در کاربرد مسئولانه هوش مصنوعی به شمار می‌رود. ایجاد سازوکارهای مؤثر برای جلوگیری از سوءاستفاده از اطلاعات کاربران، مدیریت صحیح داده‌ها و بهره‌گیری از الگوریتم‌ها در چارچوب قوانین و مقررات مالکیت معنوی از الزامات مهم حکمرانی داده در عصر هوش مصنوعی است. همچنین صیانت از حقوق تولیدکنندگان داده و محتوا و رعایت اصول مالکیت فکری در توسعه فناوری‌های هوشمند، نقش مهمی در شکل‌گیری نظامی منسجم و قابل‌اعتماد در حوزه فناوری اطلاعات ایفا می‌کند (مشارکت‌کننده شماره ۵). در کنار این موارد، شفافیت در عملکرد سامانه‌های هوش مصنوعی، تفسیرپذیری مدل‌های پیچیده و پذیرش مسئولیت در قبال تصمیم‌های مبتنی بر هوش مصنوعی ازجمله مؤلفه‌های اساسی پاسخگویی و اعتمادسازی در این حوزه محسوب می‌شوند. هنگامی که فرایندهای تصمیم‌گیری الگوریتم‌ها برای کاربران و نهادهای نظارتی قابل‌درک و بررسی

باشـد، زمینـه بـرای افزایـش اعتمـاد عمومـی بـه فناوری‌هـای هوشـمند فراهـم می‌شـود. از سـوی دیگـر، تعییـن مسـئولیت در قبـال پیامدهـای تصمیم‌هـای مبتنـی بـر هوش مصنوعـی می‌توانـد بـه ایجـاد نظامـی پاسـخگو و قابل‌پیگیـری در اسـتفاده از ایـن فنـاوری کمـک کنـد (مشـارکت‌کننده شـماره ۱۱). از منظـر توسـعه پایـدار نیـز توجـه بـه ابعـاد فرهنگـی، اجتماعـی و ارزشـی جامعـه در طراحـی و کاربـرد هوش مصنوعـی اهمیـت زیـادی دارد. حضـور و نظارت انسـان در کنـار سـامانه‌های هوشـمند، بهبـود عملکـرد الگوریتم‌هـا در چارچـوب ارزش‌هـای انسـانی و توجـه بـه پیامدهـای اجتماعـی فنـاوری ازجملـه اقداماتـی هسـتند کـه می‌تواننـد بـه توسـعه مسـئولانه و پایـدار هوش مصنوعـی کمـک کننـد. در ایـن میـان، حفـظ کرامـت و شـرافت انسـانی، انجـام آزمـون و ارزیابـی مـداوم الگوریتم‌هـا و تـلاش بـرای افزایـش دقـت و قابلیـت اعتمـاد سیسـتم‌های هوشـمند از اصـول مهـم در طراحـی اخلاق‌محـور الگوریتم‌هـا بـه شـمار می‌آینـد (مشـارکت‌کننده شـماره ۸). در بعـد عدالـت اجتماعـی نیـز ارائـه خدمـات بـدون تعصـب، توجـه بـه تنـوع و شـمول اجتماعـی و کاهـش سـوگیری در الگوریتم‌هـا از مهم‌تریـن مؤلفه‌هـای عدالـت الگوریتمـی محسـوب می‌شـوند. هـدف از ایـن رویکـرد آن اسـت کـه تصمیم‌گیری‌هـای مبتنـی بـر داده موجـب تبعیـض یـا نابرابـری در میـان گروه‌هـای مختلـف اجتماعـی نشـود. در همیـن راسـتا، پالایـش داده‌هـای معیـوب و افزایـش دقـت داده‌هـا می‌توانـد بـه بهبـود کیفیـت تصمیم‌گیری‌هـای مبتنـی بـر هوش مصنوعـی کمـک کنـد (مشـارکت‌کننده شـماره ۸). عـلاوه بـر ایـن، توسـعه هوش مصنوعـی نیازمنـد توجـه جـدی بـه سـرمایه انسـانی و ارتقـای توانمندی‌هـای علمـی جامعـه اسـت. تربیـت نیـروی انسـانی متخصـص، توسـعه مهارت‌هـای فناورانـه و اسـتفاده نوآورانـه از هوش مصنوعـی در مدیریـت اطلاعـات می‌توانـد زمینه‌سـاز پیشـرفت علمـی و فنـاوری در جامعـه باشـد. در حـوزه رسـانه و اطلاعـات نیـز هوش مصنوعـی می‌توانـد نقـش مهمـی در تشـخیص اخبـار جعلـی، تولیـد محتـوای دقیـق و معتبـر و مدیریـت صحیـح جریـان اطلاعـات ایفـا کنـد و از ایـن طریـق بـه حفـظ سـلامت فضـای رسـانه‌ای و کاهـش آسـیب‌های اجتماعـی کمـک نمایـد. درنهایـت، گسـترش هوش مصنوعـی می‌توانـد پیامدهـای گسـترده‌ای در ابعـاد فرهنگـی، اجتماعـی و اقتصـادی بـه همـراه داشـته باشـد. جلوگیـری از شـکاف‌های اجتماعـی و نسـلی، توجـه بـه تنـوع فرهنگـی در طراحـی فناوری‌هـا، مدیریـت صحیـح تحـولات ناشـی از کاربـرد گسـترده هوش مصنوعـی و پاسـخگویی بـه ابهامـات افـکار عمومـی ازجملـه اقداماتـی هسـتند کـه می‌تواننـد بـه افزایـش آگاهـی عمومـی و پذیـرش اجتماعـی ایـن فنـاوری کمـک کننـد. همچنین تشـویق مشـارکت اجتماعـی در اسـتفاده و سیاسـت‌گذاری فنـاوری و توجـه بـه نیازهـای

واقعی جامعه می‌تواند مسیر توسعه مسئولانه و اخلاق‌محور هوش مصنوعی را هموار سازد و به شکل‌گیری جامعه‌ای آگاه، عادلانه و مبتنی بر اعتماد در عصر فناوری‌های هوشمند منجر شود (مشارکت‌کننده شماره ۱۴) (جدول ۳).

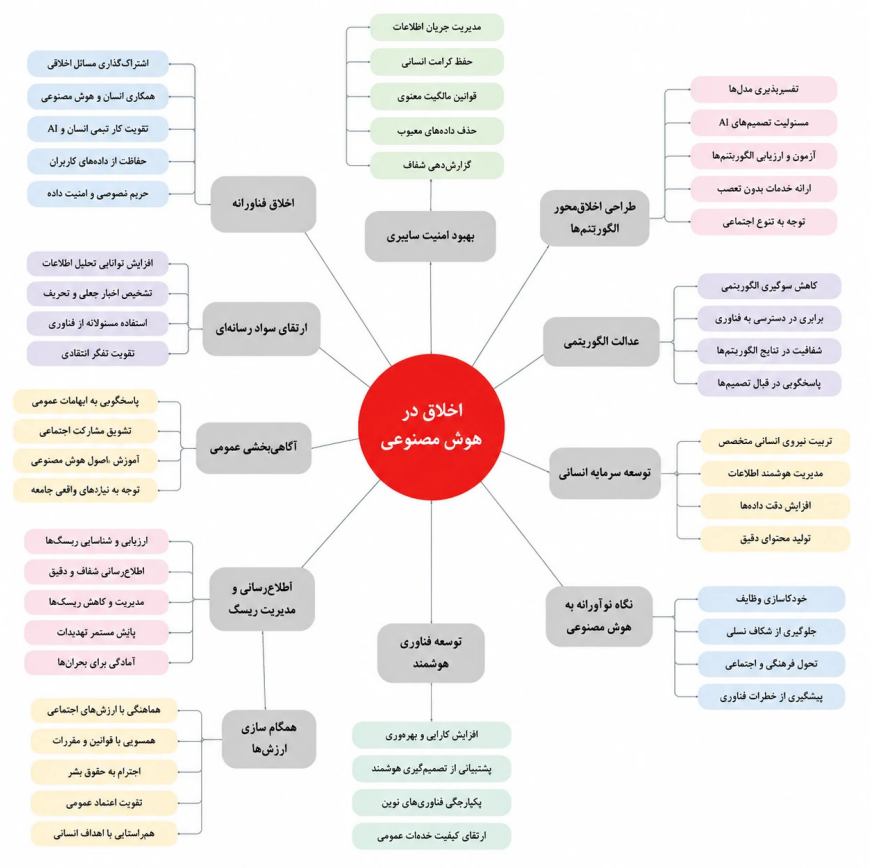
جدول ۳. اولویت مضامین اخلاق حرفه‌ای استفاده هوش مصنوعی در رسانه‌های ایران

اولویت	شاخص‌ها (بر اساس مضامین فراگیر)
۱	اخلاق فناورانه در توسعه و به‌کارگیری سامانه‌های هوش مصنوعی
۲	بهبود امنیت سایبری و حفاظت از حریم خصوصی و داده‌های کاربران
۳	حکمرانی و مدیریت دولتی در حوزه هوش مصنوعی و تنظیم مقررات مالکیت معنوی
۴	لزوم توسعه پایدار در طراحی و کاربرد فناوری‌های هوشمند
۵	طراحی اخلاق‌محور الگوریتم‌ها با تأکید بر کرامت انسانی و ارزیابی مستمر سیستم‌ها
۶	تحقق عدالت الگوریتمی و جلوگیری از تبعیض در خدمات مبتنی بر هوش مصنوعی
۷	توسعه سرمایه انسانی و تربیت نیروی متخصص در حوزه هوش مصنوعی
۸	نگاه نوآورانه به هوش مصنوعی در مدیریت اطلاعات و نظام رسانه‌ای
۹	توسعه فناوری هوشمند و بهبود کیفیت داده‌ها و عملکرد سامانه‌ها
۱۰	همگام‌سازی ارزش‌های اجتماعی و فرهنگی با تحولات ناشی از هوش مصنوعی
۱۱	اطلاع‌رسانی و مدیریت ریسک در مواجهه با چالش‌ها و تهدیدهای فناوری
۱۲	آگاهی‌بخشی عمومی و افزایش مشارکت اجتماعی در استفاده از هوش مصنوعی
۱۳	ارتقای سواد رسانه‌ای و توجه به نیازهای واقعی جامعه در توسعه فناوری‌های هوشمند

اولویت‌بندی و چینش ۱۳ مضمون فراگیر ارائه‌شده در «جدول ۳»، بر اساس تحلیل فراوانی کدها تعداد تکرار عبارات کلیدی و مضامین پایه مرتبط با هر محور در نرم‌افزار MAXQDA و همچنین میزان تأکید، اشباع مفهومی و ضریب توافق میان مشارکت‌کنندگان در مصاحبه‌های نیمه‌ساختاریافته انجام گرفته است؛ به‌گونه‌ای که مضامین دارای بالاترین بسامد تکرار و بیشترین حساسیت از دیدگاه خبرگان، در رتبه‌های نخست جدول قرار گرفته‌اند. در این راستا، توجه به اولویت‌بندی شاخص‌های اخلاق در حوزه هوش مصنوعی نشان می‌دهد که توسعه و به‌کارگیری این فناوری نیازمند شکل‌گیری چارچوبی جامع و منسجم از ملاحظات اخلاقی، حقوقی، اجتماعی و فناورانه است. در این میان، اخلاق فناورانه به‌عنوان یکی از مهم‌ترین مؤلفه‌ها مطرح می‌شود که بر طرح و تبادل مسائل اخلاقی، تقویت تعامل میان انسان و سامانه‌های هوشمند و پذیرش مسئولیت در قبال تصمیم‌های مبتنی بر الگوریتم‌ها تأکید دارد؛ امری که می‌تواند به استفاده مسئولانه از فناوری و افزایش اعتماد عمومی به سیستم‌های هوشمند منجر شود. در کنار این موضوع، امنیت سایبری و حفاظت از داده‌های کاربران

نیـز از اولویت‌هـای اساسـی بـه شـمار مـی‌رود، زیـرا بـا گسـترش کاربـرد هوش مصنوعـی، صیانـت از حریـم خصوصـی و جلوگیـری از سوءاسـتفاده از اطلاعـات شـخصی بـه یکـی از دغدغه‌هـای مهـم جوامـع تبدیـل شـده اسـت. در همیـن راسـتا، حکمرانـی مؤثـر در حـوزه هوش مصنوعـی و رعایـت قوانیـن و مقـررات مرتبـط بـا مالکیـت معنـوی و مدیریـت داده‌هـا می‌توانـد بـه تنظیـم نحـوه اسـتفاده از اطلاعـات، حمایـت از حقـوق تولیدکننـدگان محتـوا و افزایـش شـفافیت در عملکـرد سـامانه‌های هوشـمند کمـک کنـد. همچنیـن توسـعه پایـدار فناوری‌هـای هوش مصنوعـی مسـتلزم توجـه بـه ارزش‌هـای فرهنگـی، اجتماعـی و انسـانی جامعـه و پیش‌بینـی سـازوکارهای نظـارت انسـانی بـرای کنتـرل پیامدهـای احتمالـی آن اسـت. از سـوی دیگـر، طراحـی و به‌کارگیـری الگوریتم‌هـای اخلاق‌محـور بـا تأکیـد بـر حفـظ کرامـت انسـانی، ارزیابـی مـداوم عملکـرد سیسـتم‌ها و ارتقـای دقـت و قابلیـت اعتمـاد آنهـا نقـش مهمـی در بهبـود کیفیـت تصمیم‌گیری‌هـای مبتنـی بـر داده ایفـا می‌کنـد. در ایـن چارچـوب، تحقـق عدالـت الگوریتمـی نیـز اهمیـت ویـژه‌ای دارد؛ به‌گونـه‌ای کـه ارائـه خدمـات بـدون تبعیـض، توجـه بـه تنـوع اجتماعـی و کاهـش سـوگیری در داده‌هـا و الگوریتم‌هـا می‌توانـد از بـروز نابرابـری در بهره‌منـدی از فنـاوری جلوگیـری کنـد. عـلاوه بـر ایـن، توسـعه سـرمایه انسـانی و تربیـت نیـروی متخصـص در حـوزه فناوری‌هـای هوشـمند از عوامـل کلیـدی در پیشـبرد کاربـرد مسـئولانه هوش مصنوعـی بـه شـمار مـی‌رود و می‌توانـد زمینه‌سـاز ارتقـای تـوان علمـی و فناورانـه جامعـه شـود. در کنـار آن، بهره‌گیـری نوآورانـه از هوش مصنوعـی در مدیریـت اطلاعـات و نظـام رسـانه‌ای، ازجملـه در شناسـایی اخبـار جعلـی، تولیـد محتـوای دقیـق و مدیریـت جریـان اطلاعـات، می‌توانـد بـه تقویـت سـلامت فضـای رسـانه‌ای و کاهـش انتشـار اطلاعـات نادرسـت کمـک کنـد. در بعـد فناورانـه نیـز بهبـود کیفیـت داده‌هـا، پالایـش داده‌هـای معیـوب و خودکارسـازی برخـی فرایندهـا می‌توانـد موجـب افزایـش کارایـی و بهـره‌وری سیسـتم‌های هوشـمند شـود، هرچنـد گسـترش ایـن فناوری‌هـا پیامدهـای اجتماعـی و فرهنگـی متعـددی نیـز بـه همـراه دارد؛ ازایـن‌رو، همگام‌سـازی تحـولات فناورانـه باارزش‌هـای اجتماعـی و فرهنگـی، توجـه بـه تغییـرات فرهنگـی و جلوگیـری از شـکل‌گیری شـکاف‌های اجتماعـی و نسـلی از اهمیـت ویـژه‌ای برخـوردار اسـت. همچنیـن مدیریـت ریسـک‌های ناشـی از فناوری‌هـای هوشـمند، شناسـایی تهدیدهـای احتمالـی و اطلاع‌رسـانی شـفاف دربـاره پیامدهـای آنهـا می‌توانـد بـه کاهـش نگرانی‌هـای عمومـی کمـک کنـد. در ایـن میـان، افزایـش آگاهـی عمومـی، ارتقـای سـواد رسـانه‌ای و تشـویق مشـارکت اجتماعـی در فرایندهـای سیاسـت‌گذاری فنـاوری ازجملـه اقداماتـی اسـت کـه می‌توانـد زمینـه

استفاده آگاهانه و مسئولانه از هوش مصنوعی را فراهم سازد و درنهایت به شکل‌گیری اکوسیستمی اخلاق‌محور و پایدار در توسعه و کاربرد این فناوری منجر شود.



نمودار ۱. شبکه مضامین

بحث و نتیجه‌گیری

یافته‌های این پژوهش نشان می‌دهد که اخلاق حرفه‌ای در حوزه هوش مصنوعی، به‌ویژه در بستر رسانه‌های ایران، مفهومی چندبعدی و میان‌رشته‌ای است که دربرگیرنده مجموعه‌ای از اصول، شاخص‌ها و الزامات نهادی، فنی و اجتماعی است. در مقام مقایسه با پیشینه پژوهش، یافته‌های این مطالعه با دیدگاه‌های غالب در اخلاق هوش مصنوعی مبنی بر ماهیت انسان‌محور و چندبعدی فناوری همسو است (مطالعاتی

نظیر فلوریدی و همکاران، ۲۰۱۸؛ جوبین و همکاران، ۲۰۱۹ که بر شفافیت و عدالت تأکید دارند). بااین‌حال، دستاورد اصلی و نوآوری این پژوهش در گذار از مباحث انتزاعی به ارائه یک نظام شاخص‌مند و بومی‌سازی‌شده برای رسانه‌های ایران است. برخلاف بسیاری از مطالعات پیشین که یا صرفاً بر جنبه‌های فنی تمرکز داشته‌اند و یا بر رویکردهای کلی حقوقی، پژوهش حاضر توانسته است با ادغام مؤلفه‌های سواد رسانه‌ای و ظرفیت‌های بومی نهادی، مدلی عملیاتی ارائه دهد که شکاف میان اصول اخلاقی نظری و عمل حرفه‌ای را پر می‌کند. این بومی‌سازی، به‌ویژه در تأکید بر عدالت اجتماعی و آگاهی‌بخشی عمومی به‌عنوان پیش‌نیازهای اعتماد، از تفاوت‌های کلیدی این پژوهش با الگوهای غربی است که اغلب بر مباحث فردگرایانه یا تکنیک‌محور تمرکز دارند. نتایج حاصل از تحلیل مصاحبه‌ها نشان داد که مهم‌ترین ابعاد اخلاق حرفه‌ای در کاربرد هوش مصنوعی در رسانه‌ها شامل اشتراک‌گذاری چالش‌های اخلاقی میان پژوهشگران و سیاست‌گذاران، حفاظت از داده‌ها و حریم خصوصی کاربران، رعایت قوانین مالکیت معنوی، مسئولیت‌پذیری در قبال تصمیم‌های مبتنی بر الگوریتم و تقویت همکاری میان انسان و سامانه‌های هوشمند است. این ابعاد بیانگر آن است که اخلاق حرفه‌ای در هوش مصنوعی صرفاً به جنبه‌های فنی محدود نمی‌شود، بلکه به‌طور مستقیم با ساختارهای اجتماعی، حقوقی و فرهنگی جامعه نیز ارتباط دارد. از سوی دیگر، تأکید بر شفافیت در گزارش‌دهی، رفتار عادلانه در ارائه خدمات، تربیت نیروی انسانی متخصص و توانایی تشخیص اخبار جعلی نشان می‌دهد که رسانه‌ها در عصر هوش مصنوعی نیازمند بازتعریف مسئولیت‌های حرفه‌ای خود هستند. در سطح مؤلفه‌ها نیز یافته‌های پژوهش نشان داد که طراحی الگوریتم‌های اخلاق‌محور، حفظ حریم خصوصی، جلوگیری از تبعیض، شفافیت و پاسخگویی در عملکرد سامانه‌ها و تولید محتوای دقیق و معتبر از مهم‌ترین الزامات اخلاقی در توسعه و کاربرد هوش مصنوعی به شمار می‌روند. این مؤلفه‌ها درواقع زیرساخت‌های لازم برای شکل‌گیری اعتماد عمومی نسبت به فناوری‌های هوشمند هستند. زمانی که الگوریتم‌ها بر اساس اصول عدالت، شفافیت و مسئولیت‌پذیری طراحی شوند، کاربران و جامعه نسبت به نتایج و پیامدهای آنها اطمینان بیشتری خواهند داشت. در مقابل، نادیده‌گرفتن این اصول می‌تواند منجر به گسترش بی‌اعتمادی، انتشار اطلاعات نادرست و تشدید شکاف‌های اجتماعی شود. ازاین‌رو، توجه به مؤلفه‌های اخلاقی در طراحی و به‌کارگیری سامانه‌های هوش مصنوعی نه‌تنها یک الزام حرفه‌ای

بلکه پیش‌شرط توسعه پایدار این فناوری محسوب می‌شود. در بخش شاخص‌ها نیز نتایج نشان داد که رعایت اصول اخلاقی، توجه به عدالت اجتماعی، بهبود امنیت سایبری و در نظر گرفتن ابعاد فرهنگی و اجتماعی جامعه از مهم‌ترین شاخص‌های اخلاق حرفه‌ای در حوزه هوش مصنوعی هستند. این شاخص‌ها نشان می‌دهند که توسعه فناوری‌های هوشمند باید در چارچوب ارزش‌های فرهنگی و اجتماعی جامعه صورت گیرد و هم‌زمان با پیشرفت‌های فنی، به پیامدهای اجتماعی و فرهنگی آن نیز توجه شود. درواقع، نادیده گرفتن زمینه‌های فرهنگی و اجتماعی می‌تواند موجب بروز چالش‌هایی همچون تبعیض الگوریتمی، گسترش اطلاعات نادرست و کاهش اعتماد عمومی شود؛ بنابراین، تدوین و اجرای چارچوب‌های اخلاقی و مقرراتی در سطح ملی می‌تواند نقش مهمی در مدیریت این چالش‌ها و هدایت توسعه هوش مصنوعی در مسیر منافع عمومی ایفا کند. از سوی دیگر، اولویت‌بندی معیارهای اخلاق حرفه‌ای در این پژوهش نشان داد که رعایت اصول اخلاقی، تقویت امنیت سایبری، ارتقای سواد رسانه‌ای و توجه به عدالت اجتماعی از مهم‌ترین اولویت‌ها در کاربرد هوش مصنوعی در رسانه‌های ایران محسوب می‌شوند. این اولویت‌ها بیانگر آن است که توسعه فناوری‌های نوین بدون توجه به ابعاد اخلاقی می‌تواند پیامدهای اجتماعی و فرهنگی گسترده‌ای به همراه داشته باشد. در مقابل، ایجاد یک چارچوب اخلاقی منسجم می‌تواند به کاهش خطرات احتمالی، افزایش اعتماد عمومی و بهره‌گیری مؤثرتر از ظرفیت‌های هوش مصنوعی در رسانه‌ها کمک کند. درمجموع می‌توان نتیجه گرفت که اخلاق حرفه‌ای در هوش مصنوعی یکی از ارکان اساسی در توسعه و بهره‌برداری مسئولانه از این فناوری در رسانه‌ها است. رعایت اصولی مانند شفافیت، عدالت، حفاظت از حریم خصوصی، مسئولیت‌پذیری، امنیت اطلاعات و مشارکت ذی‌نفعان می‌تواند زمینه‌ساز شکل‌گیری اعتماد عمومی و استفاده پایدار از هوش مصنوعی در جامعه باشد. همچنین توجه به آموزش و توسعه سرمایه انسانی، ارتقای سواد رسانه‌ای و تدوین سیاست‌ها و مقررات مناسب می‌تواند به مدیریت چالش‌های اخلاقی این فناوری کمک کند؛ بنابراین پیشنهاد می‌شود که سیاست‌گذاران، پژوهشگران و فعالان حوزه رسانه با همکاری یکدیگر به تدوین چارچوب‌های اخلاقی و راهبردهای اجرایی در زمینهٔ کاربرد مسئولانه هوش مصنوعی بپردازند تا از ظرفیت‌های این فناوری در جهت توسعه پایدار و ارتقای منافع عمومی بهره‌برداری شود.

پیشنهادها

با توجه به یافته‌های پژوهش و اهمیت توجه به اخلاق حرفه‌ای در توسعه و به‌کارگیری فناوری‌های مبتنی بر هوش مصنوعی، ارائه مجموعه‌ای از پیشنهادهای کاربردی و پژوهشی می‌تواند به ایجاد چارچوبی منسجم برای استفاده مسئولانه از این فناوری کمک کند.

پیشنهادهای کاربردی

نخست، لازم است نهادهای سیاست‌گذار و قانون‌گذار با تدوین و تصویب چارچوب‌های اخلاقی و مقرراتی مشخص، زمینه استفاده مسئولانه از هوش مصنوعی را فراهم سازند. تدوین استانداردهای اخلاقی در طراحی، توسعه و استفاده از سامانه‌های هوش مصنوعی می‌تواند به شفافیت، پاسخگویی و کاهش چالش‌های اخلاقی در این حوزه کمک کند.

دوم، سازمان‌ها و نهادهای فعال در حوزه فناوری باید اصول اخلاق حرفه‌ای را در فرایند طراحی و توسعه سامانه‌های هوش مصنوعی مدنظر قرار دهند. طراحی الگوریتم‌های شفاف و قابل تفسیر، کاهش سوگیری‌های احتمالی در داده‌ها و الگوریتم‌ها، رعایت عدالت و جلوگیری از تبعیض در تصمیم‌های مبتنی بر هوش مصنوعی ازجمله اقداماتی است که می‌تواند به توسعه اخلاق‌محور این فناوری کمک کند.

سوم، توجه به حفاظت از داده‌ها و حریم خصوصی کاربران باید به‌عنوان یکی از مهم‌ترین اصول اخلاق حرفه‌ای در استفاده از هوش مصنوعی موردتوجه قرار گیرد. ایجاد سازوکارهای امنیتی مناسب برای مدیریت داده‌ها، جلوگیری از دسترسی غیرمجاز به اطلاعات و رعایت حقوق کاربران در جمع‌آوری و استفاده از داده‌ها می‌تواند نقش مهمی در افزایش اعتماد عمومی داشته باشد.

چهارم، توسعه سرمایه انسانی و آموزش تخصصی در حوزه اخلاق هوش مصنوعی ضروری است. برگزاری دوره‌های آموزشی، کارگاه‌ها و برنامه‌های آموزشی برای متخصصان فناوری، مدیران و کاربران می‌تواند به افزایش آگاهی نسبت به چالش‌ها و الزامات اخلاقی در این حوزه کمک کند.

پنجم، ایجاد سازوکارهای نظارتی برای ارزیابی عملکرد سامانه‌های هوش مصنوعی و میزان رعایت اصول اخلاقی در آنها ضروری است. تشکیل کمیته‌های تخصصی اخلاق در حوزه فناوری و ارزیابی مستمر الگوریتم‌ها می‌تواند به شناسایی و اصلاح

چالش‌های اخلاقی کمک کند.

ششم، مشارکت ذی‌نفعان مختلف ازجمله پژوهشگران، متخصصان فناوری، سیاست‌گذاران و جامعه در فرایند توسعه و به‌کارگیری هوش مصنوعی باید موردتوجه قرار گیرد. این مشارکت می‌تواند به شناسایی بهتر چالش‌های اخلاقی و ارائه راهکارهای مناسب برای مدیریت آنها کمک کند.

هفتم، توجه به پیامدهای اجتماعی و فرهنگی توسعه هوش مصنوعی از دیگر الزامات اخلاق حرفه‌ای در این حوزه است. در این راستا لازم است توسعه فناوری‌های هوشمند در چارچوب ارزش‌های انسانی، فرهنگی و اجتماعی جامعه انجام شود تا از بروز پیامدهای منفی و شکاف‌های اجتماعی جلوگیری گردد.

پیشنهادهای پژوهشی

در حوزه پژوهش نیز توجه به ابعاد مختلف اخلاق حرفه‌ای در هوش مصنوعی می‌تواند به توسعه دانش و سیاست‌گذاری‌های مناسب در این زمینه کمک کند. ازجمله موضوعات پیشنهادی برای پژوهش‌های آینده می‌توان به بررسی چارچوب‌های اخلاقی در توسعه و کاربرد هوش مصنوعی، تحلیل چالش‌های اخلاقی ناشی از استفاده از الگوریتم‌ها و داده‌های کلان و مطالعه نقش شفافیت و پاسخگویی در افزایش اعتماد عمومی به سامانه‌های هوش مصنوعی اشاره کرد.

همچنین انجام پژوهش‌هایی درباره سوگیری الگوریتمی و راهکارهای کاهش آن، بررسی چالش‌های مرتبط با حریم خصوصی و امنیت داده‌ها در سامانه‌های هوش مصنوعی و مطالعه تأثیرات اجتماعی و فرهنگی استفاده از این فناوری بر جوامع انسانی از دیگر موضوعات مهم برای مطالعات آینده محسوب می‌شوند. افزون بر این، طراحی مدل‌ها و شاخص‌هایی برای ارزیابی میزان رعایت اصول اخلاق حرفه‌ای در سامانه‌های هوش مصنوعی می‌تواند به توسعه چارچوب‌های علمی و عملی در این حوزه کمک کند.

منابع و مآخذ

امین بیدختی، علی‌اکبر و ابراهیم مردانی (۱۳۹۴). ارتباط بین سلامت روان و رعایت اخلاق حرفه‌ای در کارکنان پرستاری، فصلنامه اخلاق پزشکی، ۳۱(۹)، ۵۰ـ۷۳.

حسینی، نجمه و عفت عباسی (۱۳۹۱). راهکارهای نهادینه کردن اخلاق حرفه‌ای در سازمان‌ها از دیدگاه آموزه‌های دینی اسلام و پژوهش‌های مدیریتی، ۱(۲)، ۱۲۹ـ۱۵۴.

خجیر، یوسف (۱۴۰۳). فناوری‌های دور و نزدیک؛ اندیشه صاحب‌نظران حوزه فناوری‌های نوین، تهران: انتشارات موجک.

خجیر، یوسف و مهسا خلعت‌آباد (۱۴۰۵). تدوین معیارهای اخلاق حرفه‌ای استفاده از هوش مصنوعی در رسانه‌های ایران، مجلس و راهبرد، ۳۳(۱۲۶)، ۱۰۱ـ۱۳۴.

https://doi.org/10.22034/mr.2025.17136.5892

راهدارپور، جواد؛ گرگانی، فرهاد؛ صیاد، محمد؛ گزمه، میثم و مهدی نیک‌فرجام نوری (۱۴۰۳). نقش فناوری‌های هوش مصنوعی در توسعه مدیریت منابع انسانی سازمان‌های دولتی ایران، فصلنامه رویکردهای پژوهشی نو در علوم مدیریت، ۱۳ (۳۳)، ۱ـ۱۲.

رحمانی، سمانه؛ ذوالفقاری، ایرج و منوچهر سماوات (۱۴۰۳). مدیریت استراتژیک اخلاق در عصر هوش مصنوعی، فصلنامه مدیریت استراتژیک هوشمند، ۳ (۴)، ۱۰۵ـ۱۲۰.
https://doi.org/bumara.3.2.15564.35879841154

صلواتی، عادل؛ رحمانی، محمود و واحد رزمان (۱۳۹۲). رابطه اخلاق حرفه‌ای با مدیریت دانش، فصلنامه اخلاق در علوم فناوری، ۴(۸)، ۱ـ۱۰.

فادیکانی، سمیه و حمید پارسانیا (۱۴۰۲). مروری نظام‌مند بر دلالت‌های اخلاقی استفاده از هوش مصنوعی در فناوری‌های دیجیتال و نسبت آن بااخلاق شکوفایی، فصلنامه علمی راهبردهای اجتماعی و فرهنگی، ۱۲(۳)، ۷۷۱ـ۷۹۸. https://doi.org/10.22034/scs.2022.160772

قلعه‌نوئی و ابوذر گودرز (۱۴۰۴). تأثیر هوش مصنوعی بر حق حریم خصوصی از منظر حقوق بشر، قانون یار، ۹ (۳۴)، ۱۱۹ـ۱۴۳. شناسه سیویلیکا: ۲۳۷۳۲۷۵

کرمپوریان، آرزو؛ ایمنی، بهزاد و پرویز ترک زبان (۱۳۹۲). بررسی نگرش به اخلاق حرفه‌ای در اعضاء هیئت‌علمی دانشکدهٔ دندانپزشکی همدان، مجله تحقیق در علوم دندانپزشکی، ۱(۱۰)، ۴۳ـ۵۱.

Akakpo, A., Gyasi, E. A., Oduro, B., & Akpabot, S. (2024). Artificial intelligence and the future of information privacy. IGI Global. https://www.igi-global.com/article/artificial-intelligence-ai-and-the-future-of-information-privacy/383050

Alzoubi, Y. I., & Mishra, A. (2025). Differential privacy and artificial intelligence: Potentials, challenges, and future avenues. EURASIP Journal on Information Security, Article 203. https://doi.org/10.1186/s13635-025-00203-9

Amin-Bidokhti, Ali Akbar, & Mardani, Ebrahim (2015). The relationship between mental health and adherence to professional ethics among nursing staff. Journal of Medical Ethics, 9(31), 50-73. **[In Persian]**

Bessen, J. (2019). Automation and jobs: When technology boosts employment. Economic Policy, 34(100), 589-626. https://doi.org/10.1093/epolic/eiaa001

Billiris, G., Gill, A., & Bandara, M. (2025). Privacy in the age of AI: A taxonomy of data risks. arXiv Preprint. https://arxiv.org/abs/2510.02357

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., … & Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. arXiv preprint arXiv:1802.07228. https://arxiv.org/abs/1802.07228

Creswell, J. W. (2003). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd ed.). Thousand Oaks, CA: Sage Publications.

Fadikani, Somayeh, & Parsania, Hamid (2023). A Systematic Review of the Ethical Implications of Using Artificial Intelligence in Digital Technologies and Its Relationship with the Ethics of Flourishing. Quarterly Journal of Social and Cultural Strategies, 12(3), 771-798. **[In Persian]**

Floridi, L. (2023). The ethics of artificial intelligence: Principles, challenges, and opportunities.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., … & Vayena, E. (2018). AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

Ghaleh-Noei, Abouzar Goudarz (2025). The Impact of Artificial Intelligence on the Right to Privacy from a Human Rights Perspective. Ghanoon-Yar, 9(34), 119-143.

Gogoll, J., & Müller, J. F. (2017). Autonomous cars: In favor of a mandatory ethics setting. Science and Engineering Ethics, 23(3), 681-700.

Gonzalez, C., et al. (2020). Artificial intelligence and environmental sustainability: Challenges and opportunities. Sustainability, 12(10), 4050-4065.

Hosseini, Najmeh, & Abbasi, Effat (2012). Strategies for institutionalizing professional ethics in organizations from the perspective of Islamic religious teachings and management research. 1(2), 129-154. **[In Persian]**

Jobin, A., Ienca, M., & Andorno, R. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Khojir, Yousef (2024). Distant and Near Technologies: Perspectives of Experts in the Field of Emerging Technologies. Tehran: Mojak Publications. **[In Persian]**

Khojir, Yousef, & Khalat-Abad, Mahsa (2024). Formulating professional ethical standards for the use of artificial intelligence in Iranian media. Quarterly Journal of Culture and Communication Studies (Parliament Research Center Journal). **[In Persian]**

King, N. (1998). Template analysis. In G. Symon & C. Cassell (Eds.), Qualitative methods and analysis in organizational research: A practical guide (pp. 118–134). SAGE Publications.

Kluge Corrêa, N., Galvão, C., Santos, J. W., Del Pino, C., Pontes Pinto, E., Barbosa, C., Massmann, D., Mambrini, R., Galvão, L., Terem, E., & de Oliveira, N. (2024). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. arXiv Preprint. https://arxiv.org/abs/2206.11922

Korobenko, D., Nikiforova, A., & Sharma, R. (2024). Towards a privacy and security-aware framework for ethical AI. arXiv Preprint. https://arxiv.org/abs/2403.08624

Krampouryan, Arezoo, Imeni, Behzad, & Torkzaban, Parviz (2013). Investigating Attitudes toward Professional Ethics among Faculty Members of the Hamadan Faculty of Dentistry. Journal of Research in Dental Sciences, 1(10), 43-51. **[In Persian]**

Lipton, Z. C. (2018). The mythos of model interpretability. Queue, 16(3), 31-57. https://doi.org/10.1145/3236386.3241340

Martin, K. D., & Zimmermann, J. (2024). Artificial intelligence and its implications for data privacy. Journal of Business Ethics. https://pubmed.ncbi.nlm.nih.gov/38954851/

McLuhan, M. (1962). The Gutenberg Galaxy: The Making of Typographic Man. University of Toronto Press.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1-38. https://doi.org/10.1016/j.artint.2018.07.007

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

Rahdarpour, Javad, Gorgani, Farhad, Sayyad, Mohammad, Gazmeh, Meysam, & Nikfarjam Nouri, Mehdi (2024). The role of artificial intelligence technologies in the development of human resource management in Iranian government organizations. Quarterly Journal of New Research Approaches in Management Sciences, 13(33), 1-12. **[In Persian]**

Rahmani, Samaneh, Zolfaghari, Iraj, & Samavat, Manouchehr (2024). Strategic management of ethics in the age of artificial intelligence. Quarterly Journal of Intelligent Strategic Management, 3(4), 105-120. **[In Persian]**

Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.

Salavati, Adel, Rahmani, Mahmoud, & Razman, Vahid (2013). The relationship between professional ethics and knowledge management. Quarterly Journal of Ethics in Science and Technology, 4(8), 1-10. **[In Persian]**

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.