An Ethical Framework for the Use of Artificial Intelligence in Reference Services of Iranian Academic Libraries

Reza Karimi
*Corresponding author, Associate Professor, Department of Knowledge and Information Science, University of Qom, Qom, Iran. Email: karimi@qom.ac.ir

Rahil Karimian
Ph.D. Student, Department of Knowledge and Information Science, Faculty of Literature and Humanities, University of Qom, Qom, Iran. Email: r.karimian@stu.qom.ac.ir

Hadiseh Heidari
Ph.D. Student, Department of Knowledge and Information Science, Faculty of Education Sciences and Psychology, Alzahra University, Tehran, Iran. Email: ha.heidari@alzahra.ac.ir

ABSTRACT

Purpose: This study aimed to identify and validate the ethical dimensions and indicators of artificial intelligence (AI) use in reference services of Iranian academic libraries, analyze the causal influence and dependence relationships among these dimensions, and develop a framework tailored to this context.

Methodology: The study is applied in purpose and employs an exploratory mixed-methods design. Preliminary indicators derived from a structured literature review were screened and validated using the Fuzzy Delphi technique with a defuzzified acceptance threshold of 0.5 (equivalent to a crisp mean of 3). The Fuzzy DEMATEL technique was then used to analyze causal relationships and determine the influence and dependence positions of the dimensions. The expert panel comprised 18 purposively selected specialists in information science and knowledge studies, artificial intelligence, academic library management, and technology ethics. Data were collected using researcher-developed questionnaires and analyzed in Microsoft Excel 2021.

Findings: Of the 24 preliminary indicators, 19 were confirmed and five were excluded. The validated indicators were categorized into six dimensions: information security (4.13), accountability (4.07), privacy (4.04), honesty (3.93), transparency (3.92), and fairness (3.20). Fuzzy DEMATEL results showed that information security, accountability, transparency, and fairness belong to the causal (influential) group of dimensions, while privacy and honesty belong to the effect (influenced) group.

Conclusion: The ethical deployment of AI in reference services requires securing information security, transparency, and accountability as infrastructural prerequisites, along with attention to honesty in responses, source traceability, reduction of system errors, and human oversight in reference interactions. The realization of trust and confidentiality depends on meeting these requirements. Levels 1 and 2 of the final framework are based on empirical findings, while Levels 3 and 4 are the result of the authors' conceptual development.

Keywords: Artificial intelligence, Reference Services, Academic Libraries, Technology Ethics, Fuzzy Delphi.

Introduction

Artificial intelligence has become a transformative technology in information systems and knowledge-based environments in recent years. Rapid advancements in machine learning, natural language processing, computer vision, and large language models have fundamentally altered how information is produced, organized, retrieved, and disseminated, creating new opportunities for information services. These developments have significantly impacted libraries as core institutions of knowledge dissemination, paving the way for intelligent systems across diverse library operations. Consequently, academic libraries have increasingly adopted AI technologies to improve service quality, accelerate response times, enhance user experience, and support information decision-making—a trend expected to expand with ongoing AI innovations (Alala, Uzoaru, & Odikwa, 2024; Gajbhiye, 2024; Cox & Tzoc, 2023; Asemi & Asemi, 2018; Azimi, Nematollahi, & Dakhsh, 2022). Reference services represent a vital domain for AI application in libraries, playing a central role in addressing users' information needs, guiding search processes, introducing reliable sources, and offering consultation. Intelligent systems enable automated responding, smart information retrieval, resource recommendation, needs analysis, and personalized services, thereby enhancing reference efficiency and quality. Nevertheless, researchers emphasize that AI does not replace librarians' professional judgment; rather, when properly deployed, it serves as a complementary tool alongside human expertise (Hussain, 2023; Yang, 2024; Alala, Uzoaru, & Odikwa, 2024; Lo, 2024).

In this study, "artificial intelligence in reference services" refers to AI-based technologies, systems, and tools utilized across all forms of academic library reference services—in-person, remote, synchronous, and asynchronous—to answer queries, guide users, support resource discovery, analyze information needs, and assist librarians in decision-making.

Methodology

This applied study employs an exploratory mixed-methods approach within fuzzy multi-criteria decision-making. First, to compile initial ethical indicators, Persian and English literature published between January 1, 2018, and May 15, 2026 (Ordibehesht 25, 1405) was reviewed across international (Scopus, Web of Science, IEEE Xplore) and domestic databases (MagIran, Noormags, IranDoc) using keywords including artificial intelligence, reference services, academic libraries, and ethical considerations. Articles addressing normative dimensions, security, transparency, and privacy in interactive library systems were selected, yielding 24 indicators across six dimensions via thematic analysis.

Second, the Fuzzy Delphi technique was applied to screen and validate these indicators. Of 25 invited experts, 18 accepted (72% response rate) and completed both Delphi rounds with 100% participation. The panel comprised Iranian specialists in information science, academic library management, AI, and technology ethics, selected via heterogeneous purposive sampling based on academic and executive records. A five-point Likert scale was quantified using triangular fuzzy numbers, and the fuzzy mean was defuzzified via the center-of-gravity method. Indicators scoring a crisp value of 3 or higher (equivalent to a 0.5 defuzzified threshold) were confirmed.

Third, the Fuzzy DEMATEL technique determined internal relationship intensities, influence, and dependence among the six ethical dimensions. A pairwise comparison questionnaire was distributed to the same 18 experts, achieving a 100% return rate. Verbal influence scales were converted into fuzzy numbers, aggregated into a fuzzy direct-relation matrix, and defuzzified via

the center-of-gravity method. Subsequently, the total-relation matrix was derived considering direct and indirect effects. Row sums (D, influence), column sums (R, dependence), prominence ($D+R$), and net causal intensity ($D-R$) were calculated for each dimension. Data were analyzed in Microsoft Excel 2021. Kendall's coefficient of concordance reached 0.742 in the second round ($\chi^2 = 307.2, p < 0.001$), confirming strong, statistically significant expert consensus.

Findings

Fuzzy Delphi results confirmed 19 of the 24 indicators and excluded five. All indicators within privacy, transparency, information security, and accountability were validated. In the fairness dimension, "algorithmic bias avoidance" and "equal access" were confirmed, whereas "respect for user diversity," "response neutrality," and "service inclusiveness" scored below 3 and were excluded. In the honesty dimension, only "accuracy of responses" was confirmed, while "citation to reliable sources" and "declaration of uncertainty" fell below the threshold. Qualitative panel feedback attributed this to two factors: conceptual overlap between "declaration of uncertainty" and "declaration of system limitations" under transparency, and the statistical nature of language models—specifically their susceptibility to hallucinating citations—prompting experts to assign source validation to human librarian oversight rather than AI automation.

Ranking dimensions by the mean scores of confirmed indicators placed information security first (4.13), followed by accountability (4.07), privacy (4.04), honesty (3.93), transparency (3.92), and fairness (3.20). Fuzzy DEMATEL findings identified transparency ($D+R = 23.844$) as the most prominent dimension in the network, followed by privacy ($D+R = 23.604$), whose negative net value ($D-R = -0.258$) indicates greater susceptibility to influence from other factors. Information security ($D-R = +0.451$) emerged as the strongest net causal driver, followed by accountability ($D-R = +0.368$). Fairness ($D-R = +0.022$) fell into the causal group, though its proximity to zero reflects a neutral boundary position. Conversely, honesty ($D-R = -0.685$) proved to be the most dependent dimension. Overall, information security, accountability, transparency, and fairness formed the causal group, while privacy and honesty comprised the effect group. Consistency checks confirmed that the sums of D and R both equaled 69.753, yielding a zero difference.

Discussion

The findings demonstrate that AI ethics in academic libraries possesses a multilayered structure requiring a conceptual distinction between "data governance principles" and "reference-specific requirements." Infrastructural indicators—such as encryption protocols, access control, information security, and organizational accountability—are essential for any library information system. However, this study's primary contribution lies in indicators unique to reference interactions. Because users present complex information problems and AI acts as a cognitive advisor and retriever, response accuracy (honesty), source traceability and limitation disclosure (transparency) to prevent hallucinations, and algorithmic bias avoidance (fairness) directly govern the validity of the user–librarian–system triad. Failing to separate these layers risks unbalanced management where technical standards are met while normative interaction needs are neglected.

These results align with Azimi et al. (2022), who mapped AI's macro-level impact across library sectors via meta-synthesis; however, the present research operationalizes that impact into distinct ethical and structural layers for academic reference services. Furthermore, the findings

corroborate Milanifar et al. (2026), confirming that smart chatbots in virtual reference rely not merely on technical infrastructure but on human, organizational, and cultural factors. Practically, implementing this framework necessitates redefining reference roles and workflows. Librarians must transition from traditional responders to information evaluation supervisors and ethical facilitators of human–AI interaction. Achieving this requires targeted AI literacy training to detect algorithmic bias and transparent protocols for verifying AI outputs before delivery. Additionally, system designers must optimize reference interfaces so patrons can clearly distinguish machine-generated reasoning from documented, citable sources.

Conclusion

Ethical AI deployment in reference services requires establishing information security, transparency, and accountability as foundational prerequisites alongside response honesty, source traceability, error reduction, and human oversight. Securing user trust and confidentiality depends on fulfilling these criteria. The resulting four-level framework integrates empirical findings with conceptual modeling. Levels 1 and 2 derive empirically from Fuzzy Delphi and DEMATEL: Level 1 encompasses causal driving dimensions (information security, accountability, transparency, fairness) that activate and steer the system, while Level 2 covers dependent effect dimensions (privacy, honesty) reliant on Level 1 performance. Levels 3 and 4 represent conceptual extensions: Level 3 models the human–interactive mechanism within the librarian–user–AI triad, and Level 4 outlines ultimate outcomes, including user trust, enhanced service quality, reduced ethical risk, and sustainable adoption.

Because causal links in Levels 3 and 4 were not empirically tested, they serve as theoretical propositions for future validation. Additionally, two limitations warrant attention: first, relying solely on information science and AI experts excludes end-user perspectives (students and faculty), limiting generalizability regarding interactive dynamics; second, the upper framework levels remain conceptually derived without statistical model fitting. Future studies should survey academic patrons to capture user expectations and empirically validate Levels 3 and 4 using structural equation modeling, fuzzy network analysis, or DANP.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CRediT Authorship Contribution Statement

Conceptualization: [Reza Karimi] – development of the research idea, formulation of the ethical framework, and theoretical conceptualization of the four-level model. Methodology: [Reza Karimi] and [Rahil Karimian] – design of the Fuzzy Delphi and Fuzzy DEMATEL procedures, questionnaire development, and analytical protocol. Data curation: [Rahil Karimian] – literature retrieval, expert panel coordination, and data management. Formal analysis: [Rahil Karimian] – execution of Fuzzy Delphi and Fuzzy DEMATEL computations in Microsoft Excel, calculation of D, R, D+R, and D−R indices, and consistency checks. Investigation: [Reza Karimi] and [Rahil Karimian] – expert recruitment, qualitative feedback coding, and interpretation of panel notes. Writing – original draft: [Hadiseh Heidari] – preparation of the initial manuscript and extended abstract. Writing – review & editing: [Hadiseh Heidari] – critical revision, language editing, and final approval. Visualization: [Rahil Karimian and Hadiseh Heidari] – preparation of tables. Supervision: [Reza Karimi] – project

oversight and quality assurance. Project administration: [Hadiseh Heidari] – coordination and scheduling.

Declaration of AI Use

The authors hereby declare that artificial intelligence tools were not used for the generation, analysis, or interpretation of the research data, nor for the drafting of substantive scientific content, the formulation of findings, or the drawing of conclusions. AI-based language tools were used solely for limited language editing purposes—including grammar checking, improving readability, and refining English phrasing—on portions of the manuscript originally drafted by the authors. All AI-assisted edits were reviewed, verified, and approved by the authors, who take full responsibility for the accuracy, integrity, and originality of the entire manuscript. No AI tool was involved in the study design, expert panel process, Fuzzy Delphi or Fuzzy DEMATEL computations, or the development of the ethical framework.

Conflict of Interest

There is no conflict of interest.

Ethical Considerations

This study was conducted in accordance with the ethical principles of research involving human participants. All expert panel members participated voluntarily and provided informed consent prior to their involvement in both Delphi rounds and the DEMATEL pairwise comparison phase. Participants were informed of the study's purpose, the voluntary nature of their participation, and their right to withdraw at any stage without consequence. In line with the confidentiality commitments of the research, the anonymity of all 18 experts was preserved throughout data collection, analysis, and reporting; no identifying information was recorded in the datasets or disclosed in the manuscript. Expert responses were aggregated before analysis so that individual judgments could not be traced. The study did not involve vulnerable populations, sensitive personal data, or intervention procedures requiring institutional ethics board approval beyond standard research ethics protocols; nevertheless, the research complied with the ethical guidelines of and with national research ethics standards. The study involved no animal subjects. No financial or non-financial incentives were offered for participation.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request. The datasets generated and analyzed during the current study—including the Fuzzy Delphi screening data, the defuzzified indicator scores, the pairwise comparison matrices, the direct-relation and total-relation matrices, and the calculated D, R, D+R, and D−R values—are not publicly deposited due to the confidentiality commitments made to the expert panel and the need to protect the anonymity of participants' individual judgments. Aggregated data necessary to reproduce the tables and figures reported in the manuscript can be shared with qualified researchers for non-commercial academic purposes, subject to a data-sharing agreement and institutional approval. Requests should be directed to the corresponding author.

Acknowledgements

The authors express their gratitude to the esteemed reviewers, whose careful reading and valuable feedback have enriched this work.

چارچوب اخلاقی برای استفاده از هوش مصنوعی در خدمات مرجع کتابخانه‌های دانشگاهی ایران

رضا کریمی
*نویسنده مسئول، دانشیار، گروه علم اطلاعات و دانش‌شناسی، دانشکده ادبیات و علوم انسانی، دانشگاه قم، قم ایران. رایانامه: karimi@qom.ac.ir

راحیل کریمیان
دانشجوی دکتری، گروه علم اطلاعات و دانش‌شناسی، دانشکده ادبیات و علوم انسانی، دانشگاه قم، قم، ایران. رایانامه: r.karimian@stu.qom.ac.ir

حدیثه حیدری
دانشجوی دکتری، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم تربیتی و روان‌شناسی، دانشگاه الزهرا، تهران، ایران. رایانامه: ha.heidari@alzahra.ac.ir

چکیده

هدف: پژوهش حاضر با هدف شناسایی و اعتبارسنجی ابعاد و شاخص‌های اخلاقی استفاده از هوش مصنوعی در خدمات مرجع کتابخانه‌های دانشگاهی ایران، تحلیل روابط اثرگذاری و اثرپذیری ابعاد و تدوین چارچوبی متناسب با این بافت انجام شد.

روش‌شناسی: پژوهش از نظر هدف کاربردی و از نظر روش، آمیخته اکتشافی است. شاخص‌های اولیه حاصل از مرور ساختاریافته متون، با تکنیک دلفی‌فازی و آستانه پذیرش دی‌فازی‌شده ۰٫۵ (معادل میانگین قطعی ۳) غربالگری و اعتبارسنجی شدند. سپس، برای تحلیل روابط علّی و تعیین جایگاه اثرگذاری و اثرپذیری ابعاد، از تکنیک دیمتل‌فازی استفاده شد. پانل خبرگان شامل ۱۸ متخصص در حوزه‌های علم اطلاعات و دانش‌شناسی، هوش مصنوعی، مدیریت کتابخانه‌های دانشگاهی و اخلاق فناوری بود که به‌صورت هدفمند انتخاب شدند. داده‌ها با پرسشنامه‌های محقق‌ساخته گردآوری و در نرم‌افزار مایکروسافت اکسل ۲۰۲۱ تحلیل شدند.

یافته‌ها: از ۲۴ شاخص اولیه، ۱۹ شاخص تأیید و پنج شاخص حذف شد. شاخص‌های تأییدشده در شش بُعد امنیت اطلاعات (۴٫۱۳)، مسئولیت‌پذیری (۴٫۰۷)، حریم خصوصی (۴٫۰۴)، صداقت (۳٫۹۳)، شفافیت (۳٫۹۲) و عدالت (۳٫۲۰) دسته‌بندی شدند. نتایج دیمتل‌فازی نشان داد امنیت اطلاعات، مسئولیت‌پذیری، شفافیت و عدالت در گروه ابعاد علّی (اثرگذار)، حریم خصوصی و صداقت در گروه ابعاد معلولی (اثرپذیر) قرار دارند.

نتیجه‌گیری: استقرار اخلاقی هوش مصنوعی در خدمات مرجع مستلزم تأمین امنیت اطلاعات، شفافیت و مسئولیت‌پذیری به‌عنوان پیش‌نیازهای زیرساختی و توجه به صداقت پاسخ‌ها، ارجاع‌پذیری منابع، کاهش خطاهای سامانه و نظارت انسانی در تعاملات مرجع است. تحقق اعتماد و محرمانگی به تأمین این الزامات وابسته است. سطوح اول و دوم چارچوب نهایی مبتنی بر یافته‌های تجربی و سطوح سوم و چهارم حاصل توسعه مفهومی نویسندگان هستند.

کلیدواژه‌ها: هوش مصنوعی، خدمات مرجع، کتابخانه‌های دانشگاهی، اخلاق فناوری، دلفی‌فازی.

مقدمه

هوش مصـــنوعی در ســـال‌های اخیر به یکی از مهم‌ترین فناوری‌های تحول‌آفرین در نظام‌های اطلاعاتی و محیط‌های دانش‌محور تبدیل شـده اسـت. پیشـرفت سـریع الگوریتم‌های یادگیری ماشـین، پردازش زبان طبیعی، بینایی ماشـین و مدل‌های زبانی بزرگ، شیوه تولید، سازماندهی، بازیابی و اشاعه اطلاعات را دستخوش تغییرات اساسی کرده و فرصت‌های تازه‌ای را برای توسـعه خدمات اطلاعاتی فراهم آورده اسـت. این تحولات، کتابخانه‌ها را نیز به‌عنوان یکی از مهم‌ترین نهادهای تولید، سـازماندهی و اشـاعه دانش تحت تأثیر قرار داده و زمینه به‌کارگیری سـامانه‌های هوشـمند را در طیف گسـترده‌ای از فعالیت‌های کتابخانه‌ای فراهم سـاخته اسـت. در سـال‌های اخیر، کتابخانه‌های دانشـگاهی با هدف ارتقای کیفیت خدمات، افزایش سـرعت پاسـخ‌گویی، بهبود تجربه کاربران و پشـتیبانی از تصـمیم‌گیری‌های اطلاعاتی، به‌طور فزاینده‌ای از فناوری‌های هوش مصـــنوعی بهره گرفته‌اند؛ روندی که انتظار می‌رود با توســـعه فناوری‌های نوین هوش مصـنوعی در سـال‌های آینده نیز ادامه یابد (آلالا و دیگران،[1] ۲۰۲۴؛ گاجبیه،[2] ۲۰۲۴؛ کاکس و تزوک،[3] ۲۰۲۳؛ عاصـمی و عاصمی، ۲۰۱۸؛ عظیمی و دیگران ، ۱۴۰۱).

یکی از مهم‌ترین حوزه‌های کاربرد هوش مصـــنوعی در کتابخانه‌ها، خدمات مرجع اســـت؛ بخشـــی که نقش محوری در پا سخ‌گویی به نیازهای اطلاعاتی کاربران، هدایت آنان در فرایند ج ست‌وجو، معرفی منابع معتبر و ارائه م شاوره اطلاعاتی ایفا می‌کند. توسـعه سـامانه‌های هوشـمند، امکان بهره‌گیری از ابزارهایی مانند بازیابی هوشـمند اطلاعات، پاسـخ‌گویی خودکار، پیشنهاد منابع، تحلیل نیازهای اطلاعاتی و خدمات شخصی‌سازی‌شده را فراهم کرده و می‌تواند کیفیت و کارایی خدمات مرجع را ارتقا دهد. بااین‌حال، اغلب پژوهشـگران تأکید کرده‌اند که هوش مصـنوعی جایگزین قضـاوت حرفه‌ای کتابداران نیست، بلکه در صورت استقرار صحیح، می‌تواند به‌عنوان ابزاری مکمل در کنار تخصص انسانی به بهبود کیفیت خدمات منجر شود (حسین، ۲۰۲۳؛ یانگ،[4] ۲۰۲۴؛ آلالا و دیگران ، ۲۰۲۴؛ لو،[5] ۲۰۲۴)

در پژوهش حاضر، منظور از «هوش مصنوعی در خدمات مرجع» مجموعه فناوری‌ها، سامانه‌ها و ابزارهای مبتنی بر هوش م صنوعی ا ست که در همه ا شکال خدمات مرجع کتابخانه‌های دان شگاهی، اعم از خدمات حضوری و غیرحضوری و نیز خدمات هم‌زمان و ناهم‌زمان، به‌منظور پا سخ‌گویی به کاربران، راهنمایی اطلاعاتی، پشتیبانی از جست‌وجوی منابع، تحلیل نیازهای اطلاعاتی و کمک به ت صمیم‌گیری کتابداران مورد ا ستفاده قرار می‌گیرند. بنابراین، دامنه این پژوهش به فناوری خاصـی مانند چت‌بات‌ها یا هوش مصـنوعی مولد محدود نیسـت، بلکه همه انواع فناوری‌های هوش مصـنوعی شـامل: چت‌بات‌ها و د ستیارهای گفتگویی (مانند ربات‌های پا سخ‌ده به پر سش‌های مرجع)، سیستم‌های تو صیه‌گر (برای پی شنهاد منابع و خدمات متناســب با نیاز کاربر)، پردازش زبان طبیعی[6](برای تحلیل پرســش‌های کاربران و جســتجوی معنایی)، یادگیری ما شین و مدل‌های زبانی بزرگ (برای د سته‌بندی، خلا صه سازی و تولید پا سخ)، سیستم‌های جستجو و بازیابی

---

1. Adewoj
2. Gajbhiye
3. Cox & Tzoc
4. Yang
5. Lo
6. Natural Language Processing (NLP)

هو شمند اطلاعات (برای بهبود دقت و سرعت بازیابی منابع)، عوامل هو شمند و اتوما سیون فرایندها (برای انجام وظایف تکراری مرجع و ارجاع) قابل ا ستفاده در خدمات مرجع را دربر می‌گیرد. در کنار فر صت‌های ایجاد شده، ا ستفاده از هوش مصـنوعی در خدمات مرجع با مجموعه‌ای از چالش‌های اخلاقی، حرفه‌ای و مدیریتی نیز همراه اسـت. موضـوعاتی مانند حفاظت از حریم خ صو صی کاربران، امنیت داده‌ها، شفافیت عملکرد سامانه‌های هو شمند، سوگیری الگوریتمی، قابلیت اعتماد به پاسخ‌های تولیدشـده، مسـئولیت‌پذیری سـازمانی و تداوم نظارت انسـانی، از مهم‌ترین مباحث مطرح در ادبیات پژوهش به شـمار می‌روند (موهانراج و دیگران[1] ۲۰۲۴؛ بیکر-برونبائر[2] ۲۰۲۱؛ یونسـکو[3] ۲۰۲۳؛ سـازمان همکاری اقت صادی و تو سعه[4] ۲۰۲۴). همچنین، برخی سامانه‌های هو شمند برای ارائه خدمات شخ صی سازی شده ممکن ا ست داده‌های کاربران را گردآوری یا پردازش کنند؛ ازاین‌رو، نحوه مدیریت داده‌ها، رعایت محرمانگی اطلاعات و شـفافیت در ا ستفاده از داده‌های کاربران، به یکی از مهم‌ترین دغدغه‌های کتابخانه‌های دان شگاهی تبدیل شده ا ست. افزون بر این، احتمال بروز سوگیری الگوریتمی، تولید پاسخ‌های نادقیق یا استنادهای غیرقابل اتکا، ضرورت نظارت حرفه‌ای کتابداران و تدوین سازوکارهای اخلاقی منا سب را بیش از پیش آ شکار ساخته ا ست (منـ ساه[5] ۲۰۲۳؛ کمی سیون اروپا[6] ۲۰۲۴). این چالش‌ها زمانی ابعاد بحرانی‌تری به خود می‌گیرند که کتابخانه‌ها بدون تدوین الزامات اخلاقی روشـن و ارزیابی پیامدهای ناخواسته فناوری، اقدام به استقرار سامانه‌های خودکار نمایند؛ امری که می‌تواند به تضعیف بعد رابطه‌محور خدمات مرجع، مخدوش شـدن بی‌طرفی دانشـگاهی و در نهایت سـلب اعتماد کاربران به محیط‌های علمی منجر گردد (ماناسـه پرابو و ونکاتش[7] ۲۰۲۵؛ آنتونیا و ایدیودی[8] ۲۰۲۴).

با وجود گسـترش پژوهش‌ها در زمینه کاربردها و پیامدهای هوش مصـنوعی در کتابخانه‌ها، مطالعات موجود عمدتاً بر معرفی قابلیت‌های فناوری، بررسی چالش‌های اخلاقی یا ارائه توصیه‌های عمومی متمرکز بوده‌اند و از تدوین چارچوب‌های کمّی، ساختاریافته و متنا سب با اقت ضائات بومی بازمانده‌اند .در نتیجه، هنوز م شخص نی ست کدام ابعاد و شاخص‌های اخلاقی برای استفاده از انواع فناوری‌های هوش مصنوعی در همه اشکال خدمات مرجع کتابخانه‌های دانشگاهی ایران از اعتبار و اجماع خبرگان برخوردارند و این ابعاد چه روابط اثرگذاری و اثرپذیری با یکدیگر دارند .به‌ویژه در نظام آموزش عالی ایران، فقدان خطم شی‌های یکپارچه اخلاقی در به‌کارگیری هوش م صنوعی در محیط‌های کتابخانه‌ای و نبود مدلی که تقدم، تأخر و ساختار شبکه روابط این مؤلفه‌ها را تبیین کند، موجب شده تا مواجهه با این فناوری به اقداماتی واکنشی، سـلیقه‌ای و منفک از یکدیگر محدود شـود . نبود چنین چارچوبی می‌تواند تصـمیم‌گیری مدیران، تدوین خطمشـی‌ها و ارزیابی اخلاقی سامانه‌های هو شمند را به مجموعه‌ای پراکنده از ا صول و تو صیه‌های کلی محدود کند که عملاً قابلیت پیاده‌سازی و سنجش در سطح عملیاتی خدمات مرجع را نخواهند داشت .بر این اساس، هدف پژوهش حاضر شناسایی و اعتبارسـنجی ابعاد و شـاخص‌های اخلاقی اسـتفاده از انواع فناوری‌های هوش مصـنوعی در همه اشـکال خدمات مرجع

---

[1]. Mohanraj et al
[2]. Baker-Brunnbauer
[3]. UNESCO
[4]. Organisation for Economic Co-operation and Development (OECD)
[5]. Mensah
[6]. European Commission
[7] Manase Prabu & Venkatesh
[8] Anthonia & Idiodi

کتابخانه‌های دانشـگاهی ایران، تحلیل روابط اثرگذاری و اثرپذیری میان ابعاد تأییدشـده بر پایه قضـاوت خبرگان و ارائه چارچوبی متناسب با این بافت است. انتظار می‌رود نتایج این پژوهش ضمن توسعه ادبیات اخلاق هوش مصنوعی در حوزه کتابداری و اطلاع‌رسانی، مبنایی برای سیاست‌گذاری، ارزیابی و به‌کارگیری مسئولانه فناوری‌های هوش مصنوعی در خدمات مرجع کتابخانه‌های دانشگاهی ایران فراهم آورد.

پیشینه پژوهش

مرور ادبیات پژوهش نشان می‌دهد که مطالعات انجام‌شده درباره هوش مصنوعی در خدمات مرجع کتابخانه‌ها را می‌توان در چهار جریان اصلی طبقه‌بندی کرد: (۱) کاربردها و ظرفیت‌های هوش مصنوعی در خدمات مرجع، (۲) آسیب‌شناسی اخلاقی و مخاطرات هنجاری، (۳) نقش کتابدار، نظارت انسانی و الزامات سازمانی، و (۴) اسناد سیاستی، الگوهای هنجاری و خلأهای روش‌شناختی. این دسته‌بندی علاوه بر ایجاد انسجام نظری و تمایز میان «چالش‌های هنجاری» و «رویکردهای سیاستی و چارچوب‌مند»، زمینه مناسبی برای تبیین خلأ پژوهش حاضر فراهم می‌کند.

۱. کاربردها و ظرفیت‌های هوش مصنوعی در خدمات مرجع

بخش قابل‌توجهی از پژوهش‌های سال‌های اخیر به بررسی قابلیت‌های هوش مصنوعی در توسعه خدمات کتابخانه‌ای، به‌ویژه خدمات مرجع، اختصاص یافته است. نتایج این مطالعات نشان می‌دهد که فناوری‌هایی مانند یادگیری ماشین، پردازش زبان طبیعی، سیستم‌های خبره، چت‌بات‌ها، مدل‌های زبانی بزرگ و سامانه‌های بازیابی هوشمند اطلاعات، امکان پاسخ‌گویی سریع‌تر، ارائه خدمات شخصی‌سازی‌شده، بهبود بازیابی اطلاعات، دسترسی شبانه‌روزی به خدمات و کاهش فعالیت‌های تکراری کتابداران را فراهم می‌کنند (عظیمی و دیگران ، ۱۴۰۱؛ تلا[1] ۲۰۲۳؛ یانگ، ۲۰۲۴؛ آلالا و دیگران، ۲۰۲۴؛ چن و چانگ[2] ۲۰۲۵؛ آدوجو و دیگران[3] ۲۰۲۶).

در این زمینه، عظیمی و دیگران (۱۴۰۱) با استفاده از روش فراترکیب نشان دادند که هوش مصنوعی تقریباً تمامی بخش‌های کتابخانه، از جمله خدمات مرجع، سازماندهی اطلاعات، مجموعه‌سازی، بازیابی اطلاعات و برنامه‌ریزی را تحت تأثیر قرار داده است . میلانی‌فر و دیگران (۱۴۰۵) نیز با تمرکز بر چت‌بات‌های هوشمند در خدمات مرجع مجازی نتیجه گرفتند که موفقیت این فناوری علاوه بر زیرساخت‌های فنی، به عوامل انسانی، سازمانی و فرهنگی نیز وابسته است.

در سطح بین‌المللی، تلا (۲۰۲۳) با مرور نظام‌مند مطالعات تجربی گزارش کرد که کتابخانه‌های دانشگاهی در کشورهای مختلف از هوش مصنوعی برای بهبود کیفیت خدمات مرجع، افزایش سرعت پاسخ‌گویی و ارتقای تجربه کاربران استفاده می‌کنند. آلالا و دیگران (۲۰۲۴) نیز نشان دادند که ربات‌های گفت‌وگو می‌توانند ضمن کاهش بار کاری کتابداران، پاسخ‌گویی هم‌زمان به کاربران متعدد را امکان‌پذیر سازند و نقش آن‌ها بیشتر به‌عنوان دستیار هوشمند کتابدار قابل تبیین است. همچنین، چن و چانگ (۲۰۲۵) با معرفی فناوری‌های RAG وGraphRAG ، کاربرد هوش مصنوعی را در توسعه خدمات مرجع چندلایه و افزایش کیفیت پاسخ‌ها نشان دادند. در واقع، مطالعات این حوزه نشان می‌دهد که تمرکز اصلی

---

[1]. Tella
[2]. Chen & Chang
[3]. Adewojo

پژوهش‌های موجود بر ظرفیت‌های فنی و مزایای عملیاتی هوش مصـــنوعی بوده اســـت. با وجود تأکید بر بهبود کیفیت خدمات، بیشتر پژوهش‌ها به فناوری‌های خاص محدود شده‌اند و کمتر به ابعاد اخلاقی و چارچوب‌های تصمیم‌گیری برای استفاده مسئولانه از انواع فناوری‌های هوش مصنوعی در خدمات مرجع پرداخته‌اند.

۲. آسیب‌شناسی اخلاقی و مخاطرات هنجاری هوش مصنوعی

هم‌زمان با توسعه کاربردهای هوش مصنوعی، پژوهشگران توجه فزاینده‌ای به پیامدهای اخلاقی این فناوری نشان داده‌اند. مرور مطالعات بیانگر آن است که مهم‌ترین چالش‌های مطرح‌شده شامل حفاظت از حریم خصوصی کاربران، امنیت داده‌ها، سوگیری الگوریتمی، نبود شفافیت، قابلیت اعتماد به پاسخ‌های تولیدشده، مسئولیت‌پذیری، آزادی فکری و رعایت عدالت اطلاعاتی است (پوراجباری و سلامی، ۱۴۰۴؛ بیکر–برونبائر، ۲۰۲۱؛ موهانراج و دیگران، ۲۰۲۴؛ یونسکو، ۲۰۲۳؛ ؛ سازمان همکاری اقتصادی و توسعه، ۲۰۲۴).

پوراجباری و سـلامی (۱۴۰۴) با تحلیل محتوای مطالعات پیشـین، پنج محور اصـلی نگرانی‌های اخلاقی را شـامل نقض حریم خصوصی، سوگیری الگوریتمی، نبود شفافیت، تهدید امنیت داده‌ها و تضعیف نقش نیروی انسانی شناسایی کردند. نتایج پژوهش آنان نشان داد که اغلب مطالعات موجود صرفاً به شناسایی چالش‌ها پرداخته‌اند و کمتر راهکارهای اجرایی برای مدیریت هم‌زمان این مخاطرات ارائه کرده‌اند.

در همین راستا، بیکر–برونبائر (۲۰۲۱) گزارش کرد که اگرچه مدیران سازمان‌های فعال در حوزه هوش مصنوعی اهمیت اخلاق را پذیرفته‌اند، اما در بسیاری از موارد فاقد دستورالعمل‌های عملی و نظام‌مند برای پیاده‌سازی اصول اخلاقی هستند. همچنین، تلا (۲۰۲۳)، یانگ (۲۰۲۴) و آدوجو و دیگران (۲۰۲۶) نشـــان دادند که محدودیت‌هایی نظیر تولید اطلاعات نادقیق، اســـتنادهای غیرواقعی، ضـــعف در درک زمینه، ســـوگیری داده‌ها و نگرانی‌های مربوط به محرمانگی اطلاعات همچنان از مهم‌ترین موانع استفاده گسترده از هوش مصنوعی در خدمات مرجع به شمار می‌روند.

پژوهش‌های موجود درباره مخاطرات اخلاقی، عمدتاً به شناسایی فهرستی از چالش‌ها پرداخته‌اند و کمتر روابط میان این مؤلفه‌ها یا میزان اهمیت نســـبی آن‌ها را بررســـی کرده‌اند. ازاین‌رو، هنوز مشـــخص نیســـت که در محیط کتابخانه‌های دانشگاهی ایران کدام ملاحظات اخلاقی از اولویت بیشتری برخوردارند و چگونه بر یکدیگر اثر می‌گذارند.

۳. نقش کتابدار، نظارت انسانی و الزامات سازمانی

جریان سـوم پژوهش‌ها بر این موضـوع تأکید دارد که هوش مصـنوعی نباید جایگزین کتابداران شـود، بلکه باید در قالب ابزاری مکمل برای ارتقای کیفیت خدمات مورد استفاده قرار گیرد. این مطالعات بر نقش کتابدار در ارزیابی کیفیت پاسخ‌ها، نظارت بر عملکرد سامانه‌های هوشمند، آموزش کاربران، مدیریت داده‌ها و حفظ ارزش‌های حرفه‌ای کتابداری تأکید دارند (یانگ، ۲۰۲۴؛ آلالا و دیگران، ۲۰۲۴؛ آدوجو و دیگران، ۲۰۲۶؛ گاجبیه، ۲۰۲۴؛ ایفلا، ۲۰۲۴).

میلانی‌فر و دیگران (۱۴۰۵) بر ضـرورت توجه هم‌زمان به عوامل انسـانی، فرهنگی و مدیریتی در اسـتقرار چت‌بات‌های هوشمند تأکید کردند.

مطالعات جدیدتر نیز نشان می‌دهد که توسعه هوش مصنوعی موجب تغییر ماهیت نقش کتابداران شده است؛ به‌گونه‌ای که کتابداران علاوه بر ارائه خدمات اطلاعاتی، وظیفه ارزیابی پاسخ‌های تولیدشده، کنترل کیفیت اطلاعات، نظارت بر عملکرد سامانه‌ها و آموزش کاربران را نیز بر عهده دارند. ازاین‌رو، آموزش حرفه‌ای، تدوین دستورالعمل‌های سازمانی، توسعه سواد هوش مصنوعی و ایجاد نظام‌های نظارتی، از الزامات استقرار مسئولانه این فناوری محسوب می‌شود.

نتایج این مطالعات نشان می‌دهد که موفقیت هوش مصنوعی بیش از آنکه به توانمندی‌های فنی وابسته باشد، به نحوه تعامل میان انسان، فناوری و ساختارهای سازمانی بستگی دارد. بااین‌حال، اغلب پژوهش‌ها این موضوع را به‌صورت توصیفی بررسی کرده‌اند و چارچوب مشخصی برای سازمان‌دهی ابعاد اخلاقی و مدیریتی ارائه نداده‌اند.

٤ .اسناد سیاستی، الگوهای هنجاری و خلأهای روش‌شناختی

چهارمین جریان پژوهش‌ها بر توسعه چارچوب‌ها، دستورالعمل‌ها و اصول اخلاقی استفاده از هوش مصنوعی متمرکز است. سازمان‌هایی یونسکو (۲۰۲۳)، سازمان همکاری اقتصادی و توسعه، (۲۰۲۴)، ایفلا (۲۰۲۴) و کمیسیون اروپا (۲۰۲۴) مجموعه‌ای از اصول شامل شفافیت، عدالت، مسئولیت‌پذیری، انسان‌محوری، حفظ حریم خصوصی، امنیت داده‌ها و پاسخ‌گویی را برای توسعه و استفاده مسئولانه از هوش مصنوعی پیشنهاد کرده‌اند.

با وجود ارزش این اسناد، بیشتر آن‌ها در سطح سیاست‌گذاری کلان تدوین شده‌اند و کمتر به شرایط خاص کتابخانه‌های دانشگاهی یا خدمات مرجع پرداخته‌اند. همچنین، اغلب مطالعات علمی نیز مؤلفه‌های اخلاقی را به‌صورت فهرستی از اصول یا توصیه‌ها ارائه کرده‌اند و به‌ندرت از روش‌های اجماع خبرگان برای اعتبارسنجی مؤلفه‌ها یا از روش‌های تحلیل ساختاری برای بررسی روابط میان آن‌ها استفاده کرده‌اند.

به بیان دیگر، در بیشتر پژوهش‌های پیشین، ابعاد اخلاقی مستقل از یکدیگر بررسی شده‌اند و میزان اثرگذاری یا اثرپذیری متقابل آن‌ها مشخص نشده است. این موضوع موجب شده است که چارچوب‌های ارائه‌شده بیشتر جنبه توصیفی داشته باشند و کمتر بتوانند به‌عنوان مبنایی برای تصمیم‌گیری، سیاست‌گذاری و ارزیابی نظام‌مند مورد استفاده قرار گیرند.

مرور مطالعات نشان می‌دهد که اگرچه اصول اخلاقی متعددی برای استفاده از هوش مصنوعی پیشنهاد شده است، اما هنوز چارچوبی بافت‌مند، اعتبارسنجی‌شده و مبتنی بر روابط ساختاری میان ابعاد اخلاقی برای خدمات مرجع کتابخانه‌های دانشگاهی وجود ندارد.

جمع‌بندی نهایی پیشینه پژوهش

بررسی مجموعه مطالعات انجام‌شده نشان می‌دهد که پژوهش‌های پیشین را می‌توان در چهار جریان اصلی شامل «کاربردها و ظرفیت‌های هوش مصنوعی»، «آسیب‌شناسی اخلاقی و مخاطرات هنجاری»، «نقش کتابدار و الزامات سازمانی»، و «اسناد سیاستی و الگوهای هنجاری» طبقه‌بندی کرد. مطالعات این حوزه به‌طور کلی نشان داده‌اند که هوش مصنوعی می‌تواند موجب ارتقای کیفیت خدمات مرجع، افزایش سرعت پاسخ‌گویی، توسعه خدمات شخصی‌سازی‌شده و بهبود دسترسی به اطلاعات شود؛ در مقابل، چالش‌هایی مانند نقض حریم خصوصی، سوگیری الگوریتمی، نبود شفافیت، کاهش قابلیت اعتماد به پاسخ‌ها و ضرورت حفظ نظارت انسانی نیز به‌طور گسترده مورد تأکید قرار گرفته است.

با وجود این، مطالعات موجود از نظر نوع کتابخانه، فناوری مورد بررسی، شکل خدمات مرجع و سطح تحلیل اخلاقی، ناهمگون هستند و اغلب به شناسایی یا توصیف مؤلفه‌های اخلاقی بسنده کرده‌اند. در بخش عمده این پژوهش‌ها، مؤلفه‌های اخلاقی به‌صورت مستقل و بدون اعتبارسنجی مبتنی بر اجماع خبرگان ارائه شده‌اند و روابط اثرگذاری و اثرپذیری میان آن‌ها نیز بررسی نشده است. افزون بر این، پژوهشی مشاهده نشد که به‌طور هم‌زمان همه انواع فناوری‌های هوش مصنوعی و همه اشکال خدمات مرجع را در بافت کتابخانه‌های دانشگاهی ایران مورد بررسی قرار داده و چارچوبی اخلاقی مبتنی بر اجماع خبرگان و تحلیل ساختاری روابط میان ابعاد ارائه کرده باشد.

بنابراین، خلأ پژوهش حاضر صرفاً فقدان یک چارچوب اخلاقی نیست، بلکه نبود چارچوبی بومی، اعتبارسنجی‌شده و برخوردار از ساختار ارتباطی میان ابعاد اخلاقی است. نوآوری این پژوهش در سه جنبه اصلی قابل تبیین است: نخست، شناسایی و بومی‌سازی ابعاد و شاخص‌های اخلاقی متناسب با خدمات مرجع کتابخانه‌های دانشگاهی ایران؛ دوم، اعتبارسنجی این ابعاد و شاخص‌ها بر پایه اجماع خبرگان با استفاده از روش دلفی‌فازی؛ و سوم، تحلیل روابط اثرگذاری و اثرپذیری مبتنی بر قضاوت خبرگان میان ابعاد تأییدشده با بهره‌گیری از روش دیمتل‌فازی. انتظار می‌رود چارچوب حاصل بتواند علاوه بر غنی‌سازی ادبیات نظری، مبنایی برای سیاست‌گذاری، ارزیابی و توسعه استفاده مسئولانه از هوش مصنوعی در خدمات مرجع کتابخانه‌های دانشگاهی ایران فراهم آورد.

سؤالات پژوهش

پژوهش حاضر درصدد پاسخ به دو پرسش زیر است:

۱) ابعاد و شاخص‌های اخلاقی استفاده از هوش مصنوعی در خدمات مرجع کدام‌اند و چگونه شناسایی و اعتبارسنجی می‌شوند؟

۲) روابط اثرگذاری و اثرپذیری میان ابعاد اخلاقی استفاده از هوش مصنوعی در خدمات مرجع چگونه تحلیل می‌شود؟

روش‌شناسی پژوهش

پژوهش حاضر از نظر هدف، کاربردی و از حیث ماهیت و روش، رویکردی آمیخته در حوزه تصمیم‌گیری چندمعیاره فازی دارد که با بهره‌گیری از تکنیک‌های دلفی‌فازی برای غربالگری و غربال شاخص‌ها و دیتمل‌فازی (راهدار و نصر، ۱۳۹۶؛ (ژو و دیگران[1] ۲۰۱۱) برای تحلیل روابط ساختاری و شبکه‌ای میان ابعاد اجرا شده است. قلمرو موضوعی پژوهش، چارچوب‌بندی اخلاقی کاربرد هوش مصنوعی در خدمات مرجع کتابخانه‌های دانشگاهی ایران است. مراحل اجرای پژوهش به شرح زیر انجام پذیرفت:

[1]. Zhou

مرحله اول: استخراج شاخص‌های اولیه از متون تخصصی

در این مرحله به‌منظور تدوین سیاهه اولیه شاخص‌های اخلاقی، منابع علمی فارسی و انگلیسی مرتبط در بازه زمانی اول ژانویه ۲۰۱۸ تا ۱۵ مه ۲۰۲۶ (اردیبهشت ۱۴۰۵) در پایگاه‌های استنادی بین‌المللی (Scopus, Web of Science, IEEE Xplore) و پایگاه‌های اطلاعاتی داخلی (مگ‌ایران، نورمگز و ایرانداک) با کلیدواژه‌هایی نظیر هوش مصنوعی، خدمات مرجع، کتابخانه‌های دانشگاهی و ملاحظات/اصول اخلاقی مورد بررسی قرار گرفتند. مقالاتی که مستقیماً به ابعاد هنجاری، امنیت اطلاعات، شفافیت و حریم خصوصی در سامانه‌های تعاملی کتابخانه پرداخته بودند گزینش شده و از تحلیل مضمونی آن‌ها، فهرستی متشکل از ۲۴ شاخص در قالب ۶ بعد اصلی جهت اعتبارسنجی در فاز دلفی استخراج گردید.

مرحله دوم: غربالگری و اعتبارسنجی شاخص‌ها با دلفی‌فازی

برای اجرای دلفی، ابتدا از ۲۵ متخصص دعوت به عمل آمد که در نهایت ۱۸ نفر دعوت را پذیرفتند (نرخ بازگشت ۷۲٪). در هر دو دور دلفی، پرسشنامه‌ها با نرخ مشارکت ۱۰۰٪ از سوی این ۱۸ نفر تکمیل گردید. ضمناً در پرسشنامه، بخش باز پاسخ برای درج یادداشت‌های کیفی و استدلال‌های خبرگان تعبیه شد تا دلایل رد یا ادغام مؤلفه‌ها ثبت گردد. برای بومی‌سازی و سنجش درجه اهمیت شاخص‌های اولیه، روش دلفی‌فازی به کار رفت. جامعه خبرگان شامل ۱۸ نفر از متخصصان حوزه علم اطلاعات و دانش‌شناسی، مدیران کتابخانه‌های دانشگاهی، متخصصان هوش مصنوعی و اخلاق فناوری در ایران بودند که با روش نمونه‌گیری هدفمند ناهمگون بر پایه سوابق علمی و اجرایی انتخاب شدند. فرایند دلفی در دو دور متوالی انجام پذیرفت؛ در دور دوم، نتایج تجمیعی و آماری دور نخست به‌صورت محرمانه در اختیار خبرگان قرار گرفت تا به اجماع دست یابند. طیف پنج‌درجه‌ای لیکرت با استفاده از مقادیر فازی مثلثی جدول (۱) کمی‌سازی شد:

جدول ۱. متغیرهای زبانی و اعداد فازی مثلثی متناظر در دلفی

امتیاز	گزینه پاسخ	عدد فازی مثلثی (U,M,L)(L,M,U)(U,M,L)
۱	بسیار کم	(۰،۰،۰٫۲۵)
۲	کم	(۰،۰٫۲۵،۰٫۵۰)
۳	متوسط	(۰٫۲۵،۰٫۵۰،۰٫۷۵)
۴	زیاد	(۰٫۵۰،۰٫۷۵،۱)
۵	بسیار زیاد	(۰٫۷۵،۱،۱)

دیدگاه هر خبره به‌صورت $\tilde{x}_i = (\bar{l}_i, \tilde{m}_i, \bar{u}_i)$ (رابطه ۱) ثبت و میانگین فازی دیدگاه ۱۸ خبره از رابطه (۲) حاصل شد:

(۲)

$$\tilde{x}_i = (\bar{l}_i, \tilde{m}_i, \bar{u}_i) = \left(\frac{1}{18}\sum_{j=1}^{18} l_{ij} , \frac{1}{18}\sum_{j=1}^{18} m_{ij} , \frac{1}{18}\sum_{j=1}^{18} u_{ij} \right)$$

فازی‌زدایی میانگین فازی با روش مرکز ثقل (رابطه ۳) و انطباق آن بر مقیاس لیکرت بر مبنای رابطه (۴) انجام پذیرفت:

(۳)

$$C_i = \frac{\bar{l}_{\mathrm{i}} + \tilde{m}_{\mathrm{i}}, + \bar{u}_{\mathrm{i}}}{3}$$

(۴)

$$S_i = 1 + 4 \times C_i$$

شاخص‌هایی که مقدار قطعی Si آن‌ها برابر با ۳ و بیشتر بود (معادل آستانه دی‌فازی‌شده ۰/۵)، مورد تأیید نهایی قرار گرفتند. مبنای انتخاب این آستانه، نقطه میانی طیف لیکرت و تضمینِ برخورداری شاخص‌های تأییدشده از حداقل اهمیتِ «متوسط» از دیدگاه خبرگان است (چنگ و لین[1] ۲۰۰۲؛ حبیبی[2] و دیگران، ۲۰۱۴)؛ در حالی که شاخص‌های دارای میانگین قطعی کمتر از ۳ حذف شدند.

مرحله سوم: مدل‌سازی ساختاری روابط با دیتمل‌فازی

به‌منظور شناسایی شدت روابط درونی، اثرگذاری و اثرپذیری میان ۶ بُعد اخلاقی، تکنیک دیتمل‌فازی اجرا گردید. پرسشنامه ماتریسی مقایسات زوجی میان دقیقاً همان ۱۸ خبره پنل دلفی توزیع شد و تمامی ۱۸ پرسشنامه با نرخ بازگشت ۱۰۰٪ گردآوری گردید. مقیاس کلامی اثرگذاری با اعداد فازی جدول ۲ سنجیده شد:

جدول ۲. طیف کلامی و اعداد فازی دیمتل

گزینه پاسخ	عدد فازی مثلثی
بدون اثر	(۰،۰،۰٫۲۵)
اثر کم	(۰،۰٫۲۵،۰٫۵۰)
اثر متوسط	(۰٫۲۵،۰٫۵۰،۰٫۷۵)
اثر زیاد	(۰٫۵۰،۰٫۷۵،۱)
اثر بسیار زیاد	(۰٫۷۵،۱،۱)

ماتریس روابط مستقیم فازی با تجمیع نظرات خبرگان از رابطه (۵) استخراج شد:

(۵)

$$\tilde{z}_{ij} = (\bar{l}_{\mathrm{i}}, \tilde{m}_{\mathrm{i}}, \bar{u}_{\mathrm{i}}) = \left(\frac{1}{18}\sum_{k=1}^{18} L_{\mathrm{ij}}^{(k)} , \frac{1}{18}\sum_{k=1}^{18} m_{\mathrm{ij}}^{(k)} , \frac{1}{18}\sum_{k=1}^{18} u_{\mathrm{ij}}^{(k)} \right)$$

[1] Cheng & Lin
[2] Habibi

روند ویرایش نشده

پس از فازی‌زدایی مقادیر با شیوه مرکز ثقل و تشکیل ماتریس قطعی روابط مستقیم $Z = [z_{ij}]6 \times 6$، نرمال‌سازی ماتریس از طریق روابط (۶) و (۷) محاسبه گردید:

(۶)

$$s = \max\left(\max_{1\le i\le 6}\sum_{j=1}^{6} z_{ij}\,,\max_{1\le j\le 6}\sum_{i=1}^{6} z_{ij}\right)$$

(۷)

$$X = \frac{Z}{s}$$

سپس ماتریس روابط کل (T) با در نظر گرفتن اثرات مستقیم و غیرمستقیم بر اساس رابطه (۸) به دست آمد:

(۸)

$$T = X(I - X)^{-1}$$

در پایان، مجموع سطرها (Di) معرف میزان اثرگذاری و مجموع ستون‌ها (Ri) معرف میزان اثرپذیری هر بُعد تعیین گردید (روابط ۹ و ۱۰):

(۹)

$$D_1 = \sum_{j=i}^{6} t_{ij}$$

(۱۰)

$$R_1 = \sum_{j=i}^{6} t_{ij}$$

و شاخص‌های برهم‌کنش/اهمیت $(D_i + R_i)$ و شدت خالص علّی $(R_i - D_i)$ برای هر بعد اخلاقی محاسبه شدند (روابط ۱۱ و ۱۲):

(۱۱)

$$P_i = D_i + R_i$$

(۱۲)

$$E_i = D_i - R_i$$

رونوشت ویرایش نشده

یافته‌های پژوهش

به‌منظور ارزیابی اعتبار قضاوت‌های خبرگانی، در ابتدا ویژگی‌های جمعیت‌شناختی اعضای پنل تخصصی مورد بررسی قرار گرفت. پنل خبرگان این پژوهش از ۱۸ نفر متخصص تشکیل شد که بر اساس معیارهایی نظیر سابقه علمی و اجرایی، تخصص در حوزه‌های علم اطلاعات و دانش‌شناسی، هوش مصنوعی، اخلاق فناوری و مدیریت کتابخانه انتخاب شدند. نتایج در جدول ۳ نشان داد که ۸ نفر ۴۴/۴ درصد) در حوزه علم اطلاعات و دانش‌شناسی، ۵ نفر (۲۷/۸درصد) در حوزه هوش مصنوعی، ۳ نفر (۱۶/۷درصد) در حوزه اخلاق فناوری و ۲ نفر (۱۱/۱درصد) در حوزه مدیریت کتابخانه تخصص داشتند. همچنین ۱۵ نفر (۸۳/۳درصد) دارای مدرک دکتری و ۳ نفر (۱۶/۷درصد) دارای مدرک کارشناسی ارشد بودند. از نظر سابقه حرفه‌ای نیز ۴ نفر (۲۲/۲درصد) دارای سابقه ۵ تا ۱۰ سال، ۹ نفر (۵۰ درصد) دارای سابقه ۱۱ تا ۲۰ سال و ۵ نفر (۲۷/۸درصد) دارای بیش از ۲۰ سال سابقه فعالیت بودند. سنجش توافق خبرگان (ضریب کندال): به‌منظور بررسی درجه اجماع و پایایی نظرات اعضای پنل ۱۸ نفره در فرایند غربالگری دلفی‌فازی، ضریب هماهنگی کندال (W) محاسبه گردید. مقدار ضریب کندال در پایان دور دوم به ۷۴۲/. با سطح معناداری آماری $X^2 = 307.2, p < 0.001$ رسید. این ضریب نشان‌دهنده وجود اجماع بالا و معنادار میان صاحب‌نظران بوده و کفایت داده‌ها برای نهایی‌سازی شاخص‌ها را تأیید می‌کند.

جدول۳. مشخصات خبرگان پژوهش

ویژگی	طبقه	فراوانی	درصد
حوزه تخصص	علم اطلاعات و دانش‌شناسی	۸	۴۴/۴
	هوش مصنوعی	۵	۲۷/۸
	اخلاق فناوری	۳	۱۶/۷
	مدیریت کتابخانه	۲	۱۱/۱
مدرک تحصیلی	دکتری	۱۵	۸۳/۳
	کارشناسی ارشد	۳	۱۶/۷
سابقه کاری	تا ۱۰ سال ۵	۴	۲۲/۲
	تا ۲۰ سال ۱۱	۹	۵۰/۰
	بیش از ۲۰ سال	۵	۲۷/۸

همان‌طور که در جدول ۴ قابل مشاهده است، پاسخ‌های خبرگان در مقیاس پنج‌درجه‌ای لیکرت، به اعداد فازی مثلثی نرمال‌شده تبدیل شدند. برای هر شاخص، میانگین کران پایین(L) ، مقدار میانی (M) و کران بالا (U) پاسخ‌های ۱۸ خبره محاسبه شد. سپس میانگین فازی به روش مرکز ثقل فازی‌زدایی و برای تفسیر به مقیاس ۱ تا ۵ تبدیل گردید. شاخص‌های دارای میانگین قطعی ۳ و بیشتر تأیید و شاخص‌های کمتر از ۳ حذف شدند.

جدول ۲. میانگین فازی، میانگین قطعی و نتیجه ارزیابی شاخص‌های چارچوب اخلاقی استفاده از هوش مصنوعی در خدمات مرجع کتابخانه‌ای

کد	بُعد	شاخص	میانگین فازی			میانگین قطعی*	نتیجه
			L	M	U		
P1	حریم خصوصی	اخذ رضایت آگاهانه برای جمع‌آوری داده‌ها	۰/۵۱۴	۰/۷۶۴	۰/۹۳۱	۳/۹۴	تأیید
P2		حداقل‌سازی جمع‌آوری داده‌های شخصی	۰/۴۵۸	۰/۷۰۸	۰/۸۸۹	۳/۷۴	تأیید
P3		محرمانگی اطلاعات کاربران	۰/۶۹۴	۰/۹۴۴	۱/۰۰۰	۴/۵۲	تأیید
P4		امکان حذف/اصلاح داده‌های کاربر	۰/۵۱۴	۰/۷۶۴	۰/۹۳۱	۳/۹۴	تأیید
T1	شفافیت	اعلام استفاده از هوش مصنوعی	۰/۵۸۳	۰/۸۳۳	۰/۹۵۸	۴/۱۷	تأیید
T2		توضیح نحوه تولید پاسخ	۰/۴۸۶	۰/۷۳۶	۰/۹۰۳	۳/۸۳	تأیید
T3		اعلام محدودیت‌های سامانه	۰/۴۸۶	۰/۷۳۶	۰/۹۰۳	۳/۸۳	تأیید
T4		قابلیت ردیابی منبع پاسخ	۰/۴۸۶	۰/۷۳۶	۰/۹۰۳	۳/۸۳	تأیید
S1	امنیت اطلاعات	کنترل دسترسی	۰/۳۴۷	۰/۵۶۹	۰/۷۷۸	۳/۲۶	تأیید
S2		رمزنگاری داده‌ها	۰/۶۸۱	۰/۹۳۱	۱/۰۰۰	۴/۴۸	تأیید
S3		پایش رخدادهای امنیتی	۰/۶۲۵	۰/۸۷۵	۰/۹۵۸	۴/۲۸	تأیید
S4		پشتیبان‌گیری ایمن	۰/۶۸۱	۰/۹۳۱	۱/۰۰۰	۴/۴۸	تأیید
A1	مسئولیت پذیری	تعیین مسئول پاسخ‌ها	۰/۵۶۹	۰/۸۰۶	۰/۹۰۳	۴/۰۴	تأیید
A2		سازوکار رسیدگی به خطا	۰/۷۵۰	۱/۰۰۰	۱/۰۰۰	۴/۶۷	تأیید
A3		نظارت کتابدار	۰/۴۸۶	۰/۷۲۲	۰/۸۷۵	۳/۷۸	تأیید
A4		مستندسازی تصمیمات	۰/۴۸۶	۰/۷۲۲	۰/۸۷۵	۳/۷۸	تأیید
F1	عدالت	عدم سوگیری الگوریتمی	۰/۳۳۳	۰/۵۵۶	۰/۷۶۴	۳/۲۰	تأیید
F2		دسترسی برابر	۰/۳۳۳	۰/۵۵۶	۰/۷۶۴	۳/۲۰	تأیید
F3		رعایت تنوع کاربران	۰/۲۰۸	۰/۴۱۷	۰/۶۲۵	۲/۶۷	رد
F4		بی‌طرفی پاسخ‌ها	۰/۰۲۸	۰/۱۳۹	۰/۳۸۹	۱/۷۴	رد
F5		فراگیری خدمات	۰/۰۵۶	۰/۱۶۷	۰/۴۱۷	۱/۸۵	رد
H1	صداقت	صحت پاسخ‌ها	۰/۵۵۶	۰/۷۶۴	۰/۸۷۵	۳/۹۳	تأیید
H2		ارجاع به منابع معتبر	۰/۰۲۸	۰/۱۳۹	۰/۳۸۹	۱/۷۴	رد
H3		اعلام عدم قطعیت	۰/۰۲۸	۰/۱۳۹	۰/۳۸۹	۱/۷۴	رد

* **میانگین قطعی**: مقدار حاصل از فازی‌زدایی میانگین فازی با روش مرکز ثقل، تبدیل‌شده به مقیاس ۱ تا ۵.

نتایج جدول ۴ نشان داد که از مجموع ۲۴ شاخص بررسی‌شده، ۱۹ شاخص تأیید و ۵ شاخص حذف شدند. همه شاخص‌های ابعاد حریم خصوصی، شفافیت، امنیت اطلاعات و مسئولیت‌پذیری تأیید شدند. در بُعد عدالت، دو شاخص «عدم سوگیری الگوریتمی» و «دسترسی برابر» تأیید شدند؛ اما شاخص‌های «رعایت تنوع کاربران»، «بی‌طرفی پاسخ‌ها» و «فراگیری خدمات» به دلیل قرار گرفتن میانگین قطعی آن‌ها در سطح کمتر از ۳ حذف شدند.

در بُعد صداقت، تنها شاخص «صحت پاسخ‌ها» تأیید شد و دو شاخص «ارجاع به منابع معتبر» و «اعلام عدم قطعیت» با میانگین قطعی ۱/۷۴ به حد نصاب نرسیدند. بررسی ارزیابی کیفی پنل خبرگان نشان می‌دهد این تصمیم متأثر از دو عامل بافتاری بوده است: نخست، هم‌پوشانی مفهومی شاخص «اعلام عدم قطعیت» با شاخص T3 («اعلام محدودیت‌های

سامانه») در بعد شفافیت که خبرگان را به سمت پرهیز از تکرار مؤلفه‌ها سوق داد؛ و دوم، ماهیت آماری مدل‌های هوش مصنوعی زبانی و آسیب‌پذیری آن‌ها در جعل منبع که سبب شد خبرگان، اعتبارسنجی منبع را نه در بطن خودکارسازی هوش مصنوعی، بلکه در سازوکار نظارتی کتابدار (A3) تعریف کنند. یادداشت روش‌شناختی: با توجه به تک‌شاخصه شدن بعد صداقت، مقایسه رتبه‌ای آن در جدول ۶ با ابعاد چندشاخصه صرفاً جنبه توصیفی میانگین دارد و باید با در نظر گرفتن این عدم توازن تفسیری قرائت گردد.

پاسخ به پرسش نخست پژوهش: شناسایی و اعتبارسنجی ابعاد و شاخص‌های اخلاقی استفاده از هوش مصنوعی در خدمات مرجع

به‌منظور پاسخ به پرسش نخست پژوهش، ابتدا ابعاد و شاخص‌های اولیه اخلاقی استفاده از هوش مصنوعی در خدمات مرجع کتابخانه‌ای از طریق مرور ساختاریافته منابع علمی استخراج شد. حاصل این مرحله، شناسایی ۲۴ شاخص در قالب شش بُعد بنیادین شامل حریم خصوصی، شفافیت، امنیت اطلاعات، مسئولیت‌پذیری، عدالت و صداقت بود. سپس، این شاخص‌ها در قالب پرسشنامه غربالگری دلفی‌فازی در اختیار اعضای پنل ۱۸ نفره خبرگان قرار گرفت تا درجه اهمیت و قابلیت پذیرش آن‌ها بر پایه بوم دانشگاهی مورد سنجش قرار گیرد. داده‌های حاصل با استفاده از منطق فازی مثلثی تحلیل و با به‌کارگیری روش مرکز ثقل، میانگین فازی هر شاخص فازی‌زدایی و به مقیاس عددی ۱ تا ۵ تبدیل گردید. بر اساس قاعده تصمیم‌گیری، شاخص‌هایی که میانگین قطعی آن‌ها برابر یا بیشتر از ۳ بود تأیید و شاخص‌های با میانگین کمتر از ۳ از چارچوب حذف شدند. نتایج جدول ۴ نشان داد که از مجموع ۲۴ شاخص پیشنهادی، ۱۹ شاخص تأیید و ۵ شاخص حذف گردیدند. همه شاخص‌های ابعاد حریم خصوصی (۴ شاخص)، شفافیت (۴ شاخص)، امنیت اطلاعات (۴ شاخص) و مسئولیت‌پذیری (۴ شاخص) به‌طور کامل حد نصاب پذیرش را کسب کردند. در مقابل، در بُعد عدالت از ۵ شاخص، ۲ شاخص تأیید و ۳ شاخص حذف شدند؛ و در بُعد صداقت از ۳ شاخص، تنها ۱ شاخص تأیید و ۲ شاخص حذف گردیدند. تحلیل تفصیلی وضعیت ابعاد شش‌گانه به شرح زیر است:

در بُعد اول حریم خصوصی، هر چهار شاخص پیشنهادی با میانگین‌های قطعی بالا به تأیید اعضای پنل رسیدند. از این‌رو، محرمانگی اطلاعات کاربران (P3) با میانگین قطعی ۴/۵۲ بالاترین امتیاز را در این بُعد به دست آورد. شاخص‌های اخذ رضایت آگاهانه برای جمع‌آوری داده‌ها (P1) و امکان حذف/اصلاح داده‌های کاربر (P4) هر دو با میانگین قطعی۹۴/ ۳ تأیید شدند. شاخص حداقل‌سازی جمع‌آوری داده‌های شخصی (P2) نیز با میانگین قطعی۷۴ /۳ در چارچوب نهایی تثبیت گردید. این یافته‌ها نشان می‌دهد از دیدگاه متخصصان، احترام به حریم شخصی کاربران و حق کنترل بر داده‌های پرس‌وجو، پیش‌شرط بنیادین مشروعیت استفاده از سامانه‌های مرجع هوشمند است.

در بُعد دوم شفافیت، تمام چهار شاخص این بُعد با اتفاق‌نظر خبرگان مورد پذیرش قرار گرفتند به عبارتیف شاخص اعلام استفاده از هوش مصنوعی (T1) با میانگین قطعی ۴/۱۷ بالاترین امتیاز را در این بُعد کسب کرد که نشان‌دهنده الزام افشای هویت غیرانسانی سامانه برای کاربر است. سه شاخص دیگر شامل توضیح نحوه تولید پاسخ (توضیح‌پذیری) (T2)، اعلام محدودیت‌های سامانه (T3) و قابلیت ردیابی منبع پاسخ (T4) هر یک با میانگین قطعی یکسان۸۳/ ۳تأیید شدند.

این امر نشان می‌دهد شفافیت کارکردی و آگاهی کاربر از دامنه توانمندی و خطاپذیری الگوریتم‌ها، از الزامات اخلاقی غیرقابل‌چشم‌پوشی است.

در بُعد سوم امنیت اطلاعات، هر چهار شاخص مرتبط با امنیت اطلاعات، نصاب پذیرش را با رتبه‌های بسیار بالا کسب کردند. دو شاخص رمزنگاری داده‌ها (S2) و پشتیبان‌گیری ایمن (S4) هر یک با میانگین قطعی ممتاز۴۸/ ۴ در صدر این بُعد قرار گرفتند. شاخص پایش رخدادهای امنیتی (S3) با میانگین قطعی۲۸/ ۴ و شاخص کنترل دسترسی (S1) با میانگین ۳/۲۶ تأیید شدند. امتیاز بالای این مؤلفه‌ها تصریح می‌کند که امنیت داده‌ها صرفاً یک متغیر فنی نیست، بلکه زیربنای اعتماد هنجاری کتابخانه در پاسداری از داده‌های علمی کاربران است.

در بُعد چهارم مسئولیت‌پذیری، هر چهار شاخص این بُعد با استقبال بالای خبرگان تأیید شدند در واقع، شاخص سازوکار رسیدگی به خطا (A2) با میانگین قطعی ۴/۶۷، بالاترین رتبه را در میان کل شاخص‌های پژوهش به خود اختصاص داد. این نتیجه نشان می‌دهد پاسخ‌گویی در برابر اشتباهات اطلاعاتی هوش مصنوعی، حیاتی‌ترین دغدغه اخلاقی خبرگان است. شاخص تعیین مسئول پاسخ‌ها (A1) با میانگین قطعی ۴/۴۰ و دو شاخص نظارت کتابدار (A3) و مستندسازی تصمیمات (A4) هر یک با میانگین قطعی ۳/۷۸ در چارچوب قرار گرفتند. این یافته‌ها مؤید اصل محوری «انسان در حلقه» است که بر اساس آن، استقرار هوش مصنوعی هرگز رافع مسئولیت مدنی و حرفه‌ای کتابداران نخواهد بود.

در بُعد پنجم عدالت، از میان پنج شاخص، دو شاخص تأیید و سه شاخص حذف شدند. از این‌رو، شاخص‌های عدم سوگیری الگوریتمی (F1) و دسترسی برابر (F2) هر دو با کسب میانگین قطعی ۳/۲۰ در چارچوب نهایی باقی ماندند. در مقابل، سه شاخص رعایت تنوع کاربران(F3: ۲/۶۷)، بی‌طرفی پاسخ‌ها (F4: ۱/۷۴) و فراگیری خدمات (F5: ۱/۸۵)، حد نصاب ۳ را کسب نکرده و حذف شدند. واکاوی یادداشت‌های کیفی خبرگان نشان می‌دهد کاهش امتیاز شاخص بی‌طرفی پاسخ‌ها ناشی از ضرورت اتکا به شواهد معتبر علمی در پاسخ‌دهی است، نه بی‌طرفی صوری. همچنین حذف رعایت تنوع کاربران و فراگیری خدمات به چالش‌های بافتاری زیرساخت‌های فناوری در کتابخانه‌های دانشگاهی ایران نسبت داده شد که خبرگان تحقق فنی آن‌ها را در فاز نخست استقرار فراتر از بلوغ زیرساختی موجود ارزیابی کردند.

در بُعد ششم صداقت، از سه شاخص پیشنهادی تنها شاخص صحت پاسخ‌ها (H1) با میانگین قطعی ۳/۹۳ به تأیید رسید؛ درحالی‌که شاخص‌های ارجاع به منابع معتبر (H2) و اعلام عدم قطعیت (H3) با میانگین قطعی ۱/۷۴ حذف گردیدند. ارزیابی کیفی نشان داد شاخص «اعلام عدم قطعیت» به دلیل هم‌پوشانی با شاخص «اعلام محدودیت‌های سامانه (T3)» حذف شد. همچنین، شاخص «ارجاع به منابع معتبر» با توجه به پدیده «توهم الگوریتمی»، توسط خبرگان نه به عنوان قابلیت مستقل هوش مصنوعی، بلکه ذیل نظارت کتابدار (A3) تعریف گردید. از این‌رو، با باقی‌ماندن تنها یک شاخص در بُعد صداقت، تأکید می‌شود بقای این بُعد صرفاً به‌منزله ارزش غایی «صحت اطلاعات» است. از منظر متدولوژیک، مقایسه رتبه‌ای این بُعد با ابعاد چندشاخصه صرفاً بیانی توصیفی از میانگین ریاضی شاخص بازمانده است و نباید به عنوان برابری وزن سازه‌ای تعبیر گردد. همان‌طور که در جدول ۵ قابل شده است، در نهایت چارچوب اخلاقی پیشنهادی برای خدمات مرجع هوشمند شامل ۱۹ شاخص نهایی در قالب ۶ بُعد به تأیید نهایی رسید.

جدول ۳. خلاصه وضعیت ابعاد بر اساس نتایج دلفی‌فازی شاخص‌ها

بُعد اخلاقی	تعداد شاخص اولیه	تعداد شاخص تأییدشده	تعداد شاخص حذف‌شده	وضعیت بُعد در چارچوب نهایی
حریم خصوصی	۴	۴	۰	باقی‌مانده در چارچوب نهایی
شفافیت	۴	۴	۰	باقی‌مانده در چارچوب نهایی
امنیت اطلاعات	۴	۴	۰	باقی‌مانده در چارچوب نهایی
مسئولیت‌پذیری	۴	۴	۰	باقی‌مانده در چارچوب نهایی
عدالت	۵	۲	۳	باقی‌مانده با شاخص‌های محدودتر
صداقت	۳	۱	۲	باقی‌مانده با شاخص‌های محدودتر
مجموع	۲۴	۱۹	۵	—

رتبه‌بندی ابعاد اخلاقی

به‌منظور تعیین اهمیت نسبی ابعاد اخلاقی تأییدشده از دیدگاه پنل خبرگان، میانگین امتیازهای قطعی شاخص‌های پذیرفته‌شده در هر بُعد محاسبه و رتبه‌بندی آن‌ها تعیین شد. نتایج این ارزیابی در جدول ۶ ارائه شده است.

جدول ٦. رتبه‌بندی ابعاد اخلاقی بر اساس میانگین شاخص‌های تأییدشده

رتبه	بُعد اخلاقی	تعداد شاخص تأییدشده	امتیاز نهایی بُعد
۱	امنیت اطلاعات	۴	۴/۱۳
۲	مسئولیت‌پذیری	۴	۴/۰۷
۳	حریم خصوصی	۴	۴/۰۴
٤	صداقت	۱	۳/۹۳
٥	شفافیت	۴	۳/۹۲
٦	عدالت	۲	۳/۲۰

نتایج مندرج در جدول ۶ نشان می‌دهد که بُعد امنیت اطلاعات با میانگین ۴/۱۳ در رتبه نخست اهمیت قرار دارد. این امر گویای آن است که از دیدگاه متخصصان، حفاظت فنی و ساختاری از داده‌ها و زیرساخت‌های اطلاعاتی در کتابخانه‌های دانشگاهی، زیربنایی‌ترین اصل برای ورود هرگونه فناوری مبتنی بر هوش مصنوعی محسوب می‌شود. بلافاصله پس از آن، بُعد مسئولیت‌پذیری با میانگین ۴/۰۷ در رتبه دوم جای گرفته است که بر ضرورت وجود سازوکارهای نظارت انسانی، مستندسازی تصمیم‌ها و جبران خطاهای احتمالی سامانه تأکید دارد.

بُعد حریم خصوصی با میانگین ۴/۰۴ رتبه سوم را به خود اختصاص داده است که بیانگر حساسیت بالای خبرگان نسبت به صیانت از هویت و داده‌های پژوهشی کاربران در محیط‌های دانشگاهی است. بُعد شفافیت با میانگین ۳/۹۲ در رتبه پنجم قرار گرفته است که بر لزوم روشن بودن فرایندها و آگاهی کاربر از هوش مصنوعی بودن تعاملات دلالت دارد. همچنین بُعد عدالت با میانگین ۳/۲۰ رتبه ششم را کسب کرده است؛ بررسی وضعیت این بُعد نشان می‌دهد اگرچه شاخص‌های بنیادین عدم سوگیری و دسترسی برابر تأیید شده‌اند، اما به دلیل شرایط زیرساختی بوم پژوهش، دامنه شمول شاخص‌های آن محدودتر ارزیابی شده است.

پاسخ به پرسش دوم پژوهش: تحلیل روابط اثرگذاری و اثرپذیری ابعاد اخلاقی

به‌منظور پاسخ به پرسش دوم پژوهش و تعیین روابط اثرگذاری و اثرپذیری میان ابعاد اخلاقی، پس از شناسایی و اعتبارسنجی ابعاد و شاخص‌ها، از تکنیک دیمتل‌فازی استفاده شد. در این مرحله، بر اساس قضاوت ۱۸ خبره، ماتریس روابط مستقیم میان شش بُعد اخلاقی شامل حریم خصوصی، شفافیت، امنیت اطلاعات، مسئولیت‌پذیری، عدالت و صداقت تشکیل شد. سپس ماتریس روابط مستقیم پس از تبدیل قضاوت‌های خبرگان به اعداد فازی مثلثی، تجمیع و نرمال‌سازی شد و در ادامه، ماتریس روابط کل به دست آمد. پس از استخراج ماتریس روابط کل، شاخص‌های D، R، D +R و D−R محاسبه شدند. در این تحلیل، D نشان‌دهنده مجموع اثرگذاری هر بُعد بر سایر ابعادو R بیانگر مجموع اثرپذیری هر بُعد از سایر ابعاد است. شاخص D +R میزان برجستگی یا مرکزیت رابطه‌ای هر بُعد را در شبکه روابط نشان می‌دهد و شاخصD−R جهت تعیین جایگاه علّی یا معلولی خالص ابعاد به کار می‌رود؛ به‌گونه‌ای که مقدار مثبت D−R نشان‌دهنده قرارگیری بُعد در گروه عوامل علّی (اثرگذار) و مقدار منفی آن نشان‌دهنده قرارگیری در گروه عوامل معلولی (اثرپذیر) است.

پیش از تفسیر نتایج، صحت الگوریتم محاسباتی دیمتل کنترل شد. در جبر ماتریسی روش دیمتل، از آنجا که مقادیر D (مجموع سطرها) و R (مجموع ستون‌ها) از یک ماتریس روابط کل واحد حاصل می‌شوند، از نظر ریاضی الزاماً باید رابطه $\sum D = \sum R$ برقرار بوده و در نتیجه $\sum(D - R) = 0$ شود. کنترل محاسبات نشان داد که مجموع مقادیر D و R هر دو برابر با ۷۵۳/۶۹ بوده و تفاضل نهایی آن‌ها صفر است. شایان ذکر است این کنترل جبری صرفاً آزمونی برای اطمینان از نبود خطای محاسباتی در روند حل ماتریسی است و دلالتی بر روایی داده‌ها ندارد. نتایج حاصل در جدول ۷ ارائه شده است.

جدول ۷. شاخص‌های اثرگذاری، اثرپذیری و وضعیت علّی ابعاد اخلاقی (دیتمل‌فازی)

بُعد اخلاقی	D اثرگذاری	R اثرپذیری	D+R برجستگی	D−R شدت خالص علّی	گروه ساختاری	رتبه برجستگی	رتبه اثرگذاری
حریم خصوصی	۱۱٫۶۷۳	۱۱٫۹۳۱	۲۳٫۶۰۴	−۰٫۲۵۸	معلولی/اثرپذیر	۲	۳
شفافیت	۱۱٫۹۷۳	۱۱٫۸۷۱	۲۳٫۸۴۴	۰٫۱۰۲	علّی/اثرگذار	۱	۱
امنیت اطلاعات	۱۱٫۸۷۲	۱۱٫۴۲۱	۲۳٫۲۹۳	۰٫۴۵۱	علّی/اثرگذار	۳	۲
مسئولیت‌پذیری	۱۱٫۵۹۹	۱۱٫۲۳۱	۲۲٫۸۳۰	۰٫۳۶۸	علّی/اثرگذار	۶	۴
عدالت	۱۱٫۴۹۹	۱۱٫۴۷۷	۲۲٫۹۷۶	۰٫۰۲۲	علّی/اثرگذار	۴	۵
صداقت	۱۱٫۱۳۷	۱۱٫۸۲۲	۲۲٫۹۵۹	−۰٫۶۸۵	معلولی/اثرپذیر	۵	۶

مجموع مقادیر D و R برابر است:

$$\sum D = \sum R = 69/753, \sum(D - R) = 0/000$$

با توجه به نزدیک بودن دامنه مقادیر D+R (بین ۲۲٫۸ تا ۲۳٫۸)، باید در نظر داشت که جایگاه ابعاد در شبکه روابط، ماهیتی طیفی و نسبی دارد و تفکیک علّی/معلولی به‌معنای قطعیت مطلق نیست. با این وجود، نتایج نشان داد که شفافیت با مقدار D+R= 23.844 برجسته‌ترین بُعد اخلاقی در شبکه روابط میان ابعاد است. این یافته بیانگر آن است که شفافیت، هم از نظر میزان تبادل در روابط متقابل و هم از نظر مرکزیت رابطه‌ای، جایگاه نخ ست را دارا ست. همچنین، شفافیت با

مقدار D-R+0.102 در گروه عوامل علّی قرار دارد؛ هرچند برآیند اثرگذاری خالص آن در قیاس با سایر عوامل علّی محدودتر است.

حریم خصوصی با مقدار برجستگی 23.604، دومین بُعد مهم در ساختار روابط اخلاقی شناخته شد. مقدار منفی D-R= 0.258 - نشان می‌دهد که این بُعد بیش از آنکه بر سایر مؤلفه‌ها اثر بگذارد، از آن‌ها اثر می‌پذیرد. از این‌رو، تحقق هنجاری حریم خصوصی در خدمات مرجع هوشمند تا حد زیادی در گرو پیاده سازی اصول امنیت اطلاعات، شفافیت و سازوکارهای مسئولیت‌پذیری است.

امنیت اطلاعات با مقدار D-R= 0.451 قوی‌ترین عامل علّی خالص در میان تمامی ابعاد است. این نتیجه نشان می‌دهد امنیت اطلاعات پایه‌ای‌ترین نقش محرک را در ساختار اخلاقی پژوهش ایفا می‌کند و ارتقای آن پیش‌نیازی اجتناب‌ناپذیر برای حفاظت از حریم خصوصی و ایجاد اعتماد و پاسخ‌گویی در سامانه خواهد بود.

مسئولیت‌پذیری نیز با کسب مقدار D-R= 0.368 دومین پیشران علّی بااهمیت تعیین شد. این یافته گویای آن است که استقرار اصل نظارت انسانی، مستندسازی تصمیم‌ها و پیش‌بینی سازوکارهای مدون جبران خطا، نقشی بنیادین در فعال‌سازی سایر الزامات اخلاقی به عهده دارد.

بُعد عدالت با مقدار D-R= 0.022 گرچه در گروه عوامل علّی دسته‌بندی می‌شود، اما نزدیکی بسیار زیاد این عدد به صفر نشان می‌دهد که این بُعد دارای حالتی خنثی و در مرز میان اثرگذاری و اثرپذیری قرار دارد؛ بنابراین عملکرد آن در سامانه، همبسته و وابسته به متغیرهای پیرامونی است.

در نهایت، بُعد صداقت با مقدار D-R= -0.685 معلولی‌ترین (اثرپذیرترین) بُعد اخلاقی در چارچوب پژوهش شناخته شد. این یافته تبیین می‌کند که صداقت و صحت خروجی‌های تولیدشده توسط هوش مصنوعی، به‌تنهایی متغیری خودمختار نیست، بلکه ماحصل کارکرد درست و هم‌افزای سایر ابعاد پیشران، به‌ویژه شفافیت سازوکارها، تدابیر امنیتی و کنترل‌های مسئولانه کتابداران است.

کنترل سازگاری محاسبات دیمتل

به‌منظور اطمینان از صحت محاسبات انجام شده در تحلیل دیمتل‌فازی، سازگاری جبری شاخص‌های استخراج شده از ماتریس روابط کل بررسی شد. در روش دیمتل، از آنجا که شاخص‌های D و R هر دو از یک ماتریس روابط کل واحد به دست می‌آیند، مجموع مقادیر D باید با مجموع مقادیر R برابر باشد. همچنین، بر اساس این رابطه، مجموع مقادیر D−R باید برابر صفر شود. کنترل محاسبات جبری ماتریس روابط کل نشان داد که مجموع مقادیر D و R برابر بوده (۶۹/۷۵۳) و تفاضل آن‌ها صفر است.

چارچوب اخلاقی پیشنهادی برای استفاده از هوش مصنوعی در خدمات مرجع کتابخانه‌ای

بر اساس نتایج حاصل از غربالگری دلفی‌فازی و تحلیل روابط ساختاری دیتمل‌فازی، چارچوب اخلاقی پژوهش در چهار سطح سازمان‌دهی شد. به‌منظور شفافیت روش‌شناختی، ساختار این چارچوب بر دو لایه متمایز استوار است: نخست، «لایه تجربی و متکی بر تحلیل داده‌ها» (سطوح ۱ و ۲) که مستقیماً برونداد محاسبات دلفی و دیتمل‌فازی است؛ و دوم، «لایه

توسعه مفهومی و تجویزی مؤلفان» (سطوح ۳ و ۴) که حاصل سنتز ادبیات تخصصی و دلالت‌های کاربردی پژوهش بوده و مدلی برای آزمون مستقیم تجربی آن‌ها در این مقاله اجرا نشده است. در لایه تجربی، نتایج دلفی‌فازی نشان داد که از میان ۲۴ شاخص اولیه، ۱۹ شاخص در قالب شش بُعد حریم خصوصی، شفافیت، امنیت اطلاعات، مسئولیت‌پذیری، عدالت و صداقت به تأیید نهایی رسیدند. اگرچه هر شش بُعد در چارچوب نهایی تثبیت شدند، توزیع شاخص‌های مورد تأیید در آن‌ها یکنواخت نبود؛ ابعاد حریم خصوصی، شفافیت، امنیت اطلاعات و مسئولیت‌پذیری هر یک با ۴ شاخص، بُعد عدالت با ۲ شاخص و بُعد صداقت با ۱ شاخص («صحت پاسخ‌ها») تأیید شدند. سپس در تحلیل ساختاری دیتمل‌فازی، جایگاه رابطه‌ای این شش بُعد تعیین گردید. از این‌رو، سطح نخست چارچوب (ابعاد زیربنایی و علّی) شامل ابعاد امنیت اطلاعات $(D-R=+0/451)$ مسئولیت‌پذیری(D-R=+0.368) ، شفافیت $(D-R=+0/102)$ و عدالت $(D-R=+0/022)$ است. این ابعاد در گروه عوامل علّی (پیشران) قرار دارند و نقش زیربنایی را در فعال‌سازی و هدایت سایر مؤلفه‌های اخلاقی ایفا می‌کنند. امنیت اطلاعات از طریق کنترل دسترسی، رمزنگاری و پایش رخدادها، بستر فنی حفاظت را می‌سازد؛ مسئولیت‌پذیری از مجرای تعریف متولی پاسخ، نظارت انسانی و مستندسازی فرایندها، پاسخ‌گویی را نهادینه می‌کند؛ شفافیت با تصریح استفاده از هوش مصنوعی و تبیین محدودیت‌ها، تعامل آگاهانه را مقدور می‌سازد؛ و عدالت از طریق پایش سوگیری‌های الگوریتمی، امکان دسترسی متوازن را پشتیبانی می‌نماید.

سطح دوم چارچوب (ابعاد پیامدی و معلولی) شامل ابعاد حریم خصوصی $(D-R=-0/258)$ و صداقت $(D-R=-0/685)$ است که در تحلیل دیمتل در گروه عوامل معلولی (اثرپذیر) جای گرفتند. این یافته تجربی نشان می‌دهد که صیانت از حریم خصوصی و دست‌یابی به صداقت و صحت پاسخ‌ها، اهدافی پیامدی هستند که به‌طور مستقل در خلأ محقق نمی‌شوند، بلکه تحقق آن‌ها وابستگی شدیدی به کارکرد هماهنگ ابعاد پیشران سطح نخست (به‌ویژه امنیت، شفافیت و مسئولیت‌پذیری) دارد. در امتداد یافته‌های تجربی فوق، سطوح سوم و چهارم چارچوب به‌عنوان «توسعه مفهومی و دلالت‌های نظری مؤلفان» جهت کاربست عملیاتی مدل پیشنهاد شده‌اند:

سطح سوم چارچوب (سازوکار تعاملی و انسانی) این سطح متکی بر بسط نظری نقش «عامل انسانی در حلقه» است و به چگونگی تعامل کتابدار، کاربر و سامانه هوش مصنوعی می‌پردازد. در این لایه مفهومی، هوش مصنوعی نقشی صرفاً کمکی و مکمل داشته و کتابدار مرجع در جایگاه ناظر، ارزیاب اعتبار محتوا و تسهیل‌گر اخلاقی قرار می‌گیرد؛ نقشی که به‌ویژه در هنگام ابهام نیاز اطلاعاتی، پیچیدگی پرسش یا حساسیت‌های هنجاری ضرورت می‌یابد.

سطح چهارم چارچوب (پیامدهای مورد انتظار): این سطح بروندادها و آثار غایی به‌کارگیری چارچوب اخلاقی را از منظر نظری صورت‌بندی می‌کند. این پیامدها شامل ارتقای اعتماد نهادی کاربران به کتابخانه، بهبود کیفیت خدمات، افزایش حس انصاف و کاهش مخاطرات اخلاقی است. شایان ذکر است روابط علّی این سطح با سطوح پیشین در این پژوهش به‌صورت تجربی آزمون نشده و به‌عنوان فرضیه‌ها و دلالت‌های نظری برای اعتبارسنجی در پژوهش‌های آتی (نظیر مدل‌سازی معادلات ساختاری) پیشنهاد می‌شود. خلاصه ساختار لایه‌ای چارچوب در جدول ۸ ارائه شده است.

جدول ۸. چارچوب اخلاقی پیشنهادی برای استفاده از هوش مصنوعی در خدمات مرجع (تفکیک یافته‌های تجربی و توسعه مفهومی)

لایه روش‌شناختی چارچوب	سطح	عنوان و ماهیت سطح	ابعاد و مؤلفه‌ها	کارکرد و منطق روش‌شناختی
لایه تجربی *(مبتنی بر داده‌های دلفی و دیتمل‌فازی)*	سطح اول	ابعاد علّی و راهبردی (پیشران)	امنیت اطلاعات، مسئولیت‌پذیری، شفافیت، عدالت	مؤلفه‌های اثرگذار خالص $D-R>0$ در ایجاد زیرساخت و قواعد عملکردی سامانه مرجع
لایه تجربی *(مبتنی بر داده‌های دلفی و دیتمل‌فازی)*	سطح دوم	ابعاد معلولی و اثرپذیر (پیامدی)	حریم خصوصی، صداقت (صحت پاسخ‌ها)	ابعاد اثرپذیر $D-R<0$ ؛ تحقق آن‌ها در گرو استقرار موفقیت‌آمیز ابعاد سطح اول است
لایه توسعه مفهومی *(استنتاج نظری مؤلفان بر پایه ادبیات)*	سطح سوم	سازوکار تعاملی و مداخله انسانی	تعامل سه‌گانه کتابدار، کاربر و سامانه هوش مصنوعی	پیشنهاد الگوهای تعاملی نظیر نظارت انسانی، کنترل اعتبار پاسخ و عدم جایگزینی کامل کتابدار
لایه توسعه مفهومی *(استنتاج نظری مؤلفان بر پایه ادبیات)*	سطح چهارم	پیامدهای غایی مورد انتظار	اعتماد کاربران، ارتقای کیفیت درک‌شده، کاهش ریسک اخلاقی و استفاده پایدار	صورت‌بندی پیامدهای کیفی و سیستمی استقرار چارچوب (پیشنهادشده جهت آزمون در مطالعات آتی)

سطوح ۱ و ۲ مبتنی بر تحلیل داده‌های حاصل از پنل خبرگان (دلفی و دیتمل‌فازی) هستند؛ در حالی که سطوح ۳ و ۴ بازتاب‌دهنده مدل‌سازی مفهومی و بسط دلالت‌های تجویزی مؤلفان بر اساس مبانی نظری بوده و فاقد آزمون تجربی مستقیم در این مقاله هستند.

بر اساس نتایج دلفی‌فازی، شاخص‌های نهایی چارچوب اخلاقی در جدول ۹ قابل مشاهده است.

جدول ۹. شاخص‌های نهایی چارچوب اخلاقی

بُعد اخلاقی	شاخص‌های تأییدشده
حریم خصوصی	اخذ رضایت آگاهانه برای جمع‌آوری داده‌ها؛ حداقل‌سازی جمع‌آوری داده‌های شخصی؛ محرمانگی اطلاعات کاربران؛ امکان حذف/اصلاح داده‌های کاربر
شفافیت	اعلام استفاده از هوش مصنوعی؛ توضیح نحوه تولید پاسخ؛ اعلام محدودیت‌های سامانه؛ قابلیت ردیابی منبع پاسخ
امنیت اطلاعات	کنترل دسترسی؛ رمزنگاری داده‌ها؛ پایش رخدادهای امنیتی؛ پشتیبان‌گیری ایمن
مسئولیت‌پذیری	تعیین مسئول پاسخ‌ها؛ سازوکار رسیدگی به خطا؛ نظارت کتابدار؛ مستندسازی تصمیمات
عدالت	عدم سوگیری الگوریتمی؛ دسترسی برابر
صداقت	صحت پاسخ‌ها

بحث و نتیجه‌گیری

تحلیل یافته‌های پژوهش حاضر نشان می‌دهد که مسئله اخلاق هوش مصنوعی در محیط کتابخانه‌های دانشگاهی، ساختاری چندلایه دارد و نیازمند تفکیک مفهومی میان «اصول حاکمیت داده» و «الزامات اختصاصی خدمات مرجع» است. بخشی از ابعاد و شاخص‌های استخراج‌شده، نظیر پروتکل‌های رمزنگاری، کنترل سطوح دسترسی، امنیت اطلاعات و چارچوب‌های پا سخ‌گویی سازمانی، ماهیتی عام و زیر ساختی دارند که برای هر گونه ا ستقرار سامانه‌های اطلاعاتی در کتابخانه ضروری محسوب می‌شوند. با این حال، ارزش افزوده این پژوهش و وجه تمایز آن با مطالعات عمومی حوزه

فناوری، در شاخص‌هایی متجلی است که منحصراً از ویژگی‌های سرشتین میز مرجع ناشی می شوند. در بافت تعاملی خدمات مرجع، کاربر با مسئله مبهم اطلاعاتی مراجعه می‌کند و هوش مصنوعی در نقش یک مشاور شناختی و بازیاب محتوا عمل می‌نماید؛ بنابراین مؤلفه‌هایی چون صحت پاسخ‌های ارائه شده (صداقت)، قابلیت ردیابی منبع پاسخ و اعلام محدودیت‌های سامانه (شفافیت) برای پیشگیری از توهم اطلاعاتی، و عدم سوگیری الگوریتمی (عدالت)، شاخص‌هایی هستند که مستقیماً اعتبار تعامل پژوهشی میان کاربر، کتابدار و سامانه را تعریف می‌کنند. عدم تفکیک این دو لایه می‌تواند منجر به اتخاذ تدابیر مدیریتی نامتوازن گردد که در آن استانداردهای فنی تأمین شده، اما نیازهای هنجاریِ تعامل اطلاعاتی نادیده انگاشته شود.

ترسیم جایگاه ابعاد در ساختار حاصل از دیمتل‌فازی، منطق مدیریتی این چارچوب را از زاویه ادراکات خبرگان تبیین می‌نماید. طبقه‌بندی ابعاد به دو دسته اثرگذار و اثرپذیر، بازتاب‌دهنده اولویت‌های ساختاری و ترتیبات پیش‌نیازی در پیاده سازی سیستم است و نباید به‌مثابه اثبات روابط علّی قطعی و مکانیکی قلمداد شود. قرارگیری امنیت، شفافیت و مسئولیت‌پذیری در جایگاه ابعاد پیشران حاکی از آن است که از دیدگاه متخصصان، مقولاتی مانند محرمانگی اطلاعات پژوهشی مراجعان و صداقت در محتوای پاسخ‌ها، به‌صورت مستقل یا صرفاً از طریق توصیه‌های اخلاقی حاصل نمی‌شوند؛ بلکه تحقق آن‌ها پیامد مستقیم پیاده‌سازی سازوکارهای امنیتی، قابلیت ممیزی الگوریتم‌ها و تدوین مسئولیت‌های حقوقی مشخص است. این یافته با اصول مندرج در اسناد یونسکو (۲۰۲۳)، سازمان همکاری اقتصادی و توسعه (۲۰۲۴) و ایفلا (۲۰۲۴) — نظیر شفافیت، مسئولیت‌پذیری، انسان‌محوری و حفظ حریم خصوصی — در سطح «انتخاب اصول» همسو است، اما ناهمسویی در «سازوکار تحقق» و «سطح تحلیل» است: آن اسناد اصول را در سطح سیاست‌گذاری کلان و به صورت فهرستی از توصیه‌ها ارائه می‌کنند و روابط اثرگذاری/اثرپذیری میان ابعاد را تبیین نمی‌کنند، حال آنکه پژوهش حاضر نشان می‌دهد امنیت، شفافیت و مسئولیت‌پذیری ابعاد پیشران‌اند و تحقق آن‌ها نه صرفاً از توصیه‌های اخلاقی، بلکه از سازوکارهای نهادی و فنی حاصل می‌شود. بنابراین، ارتقای سلامت تعاملی در خدمات مرجع مستلزم تمرکز برنامه‌ریزان بر اصلاح و تقویت لایه‌های اثرگذار سامانه است تا بسترهای لازم برای تجلی ابعاد اثرپذیر فراهم گردد؛ توصیه‌ای عملیاتی که از دل تحلیل ساختاری برمی‌آید و در اسناد سیاستی کلان به‌صورت مشخص ارائه نشده است.

تطبیق این نتایج با ادبیات تجربی موجود، تفاوت‌های بافتاری و روش شناختی این پژوهش را آشکار می سازد. درحالی‌که پژوهش‌هایی نظیر تلا (۲۰۲۳) چالش‌های مربوط به حریم خصوصی، شفافیت و سوگیری الگوریتمی را در قالب موانع مجزا و بدون ارتباط ساختاری گزارش کرده‌اند، یافته‌های مطالعه حاضر نشان داد که این مفاهیم در یک شبکه به‌هم‌پیوسته قرار دارند؛ به‌طوری‌که حریم خصوصی بیش از آنکه یک متغیر مستقل باشد، تابعی ساختاریافته از سامانه‌های امنیتی و رویه‌های شفاف گزارش‌دهی است. همچنین، ضرورت حفظ نظارت انسانی و استقرار الگوی انسان در حلقه، در توافق کامل با مبانی نظری یانگ (۲۰۲۴) و آلالا و دیگران (۲۰۲۴) قرار دارد. با این تفاوت که در زیست‌بوم کتابخانه‌های دانشگاهی ایران، به دلیل عدم تدوین اسناد سیاستی بالادستی و استانداردهای حرفه‌ای بومی، حضور کتابدار نه صرفاً یک ابزار پایش کیفی، بلکه اهرم اصلی جبران خلأهای سیاستی و مهار ریسک‌های حقوقی است. از سوی دیگر، تأکید خبرگان بر شاخص قابلیت ردیابی استنادی در بعد صداقت، با معماری‌های فنی متأخر مانند مدل‌های بازیابی‌محور در مطالعه چن و چانگ (۲۰۲۵) هم‌پوشانی دارد و نشان می‌دهد که الزامات اخلاقیِ استخراج شده، دارای مابه‌ازای فنی و عملیاتی دقیق

در سامانه‌های مدرن هستند. این یافته‌ها با نتایج عظیمی و دیگران (۱۴۰۱) که با روش فراترکیب نشان دادند هوش مصنوعی تقریباً تمامی بخش‌های کتابخانه، از جمله خدمات مرجع، سازماندهی اطلاعات، مجموعه‌سازی، بازیابی اطلاعات و برنامه‌ریزی را تحت تأثیر قرار داده است، همسو هستند؛ با این تفاوت که آنان گستره تأثیر را در سطح کلان و توصیفی ترسیم کرده‌اند، حال آنکه پژوهش حاضر همان تأثیر را در بافت اختصاصی خدمات مرجع دانشگاهی به لایه‌های اخلاقی و ساختاری تفکیک و عملیاتی می سازد. همچنین، یافته حا ضر با نتیجه میلانی‌فر و دیگران (۱۴۰۵) که با تمرکز بر چت‌بات‌های هوشمند در خدمات مرجع مجازی نشان دادند موفقیت این فناوری علاوه بر زیرساخت‌های فنی، به عوامل انسانی، سازمانی و فرهنگی نیز وابسته است، هم‌سویی نظری آشکاری دارد؛ زیرا هر دو بر وابستگی چندلایه به عوامل فنی و انسانی— سازمانی تأکید می‌کنند. با این حال، تفاوت در شدت و جایگاه عامل انسانی است: از سویی دیگر، ان‌ها عوامل انسانی را در کنار عوامل فنی و سازمانی به‌صورت هم‌ارز مطرح می‌کنند، اما پژوهش حاضر نقش کتابدار را در بافت ایران—به‌دلیل خلأ استناد سیاستی بالادستی و استانداردهای حرفه‌ای بومی—فراتر برده و آن را اهرم اصلی جبران خلأهای سیاستی و مهار ریسک‌های حقوقی می‌داند.

از حیث دلالت‌های کاربردی، استقرار این چارچوب نیازمند بازتعریف نقش‌های شغلی و بازنگری در طراحی فرایندهای خدمات مرجع است. کتابداران مرجع باید از موقعیت پاسخ‌دهنده سنتی به ناظران ارزیابی اطلاعات و تسهیل‌گران اخلاقیِ تعامل انسان و هوش مصنوعی تغییر موضع دهند. این امر مستلزم طراحی دوره‌های آموزش سواد هوش مصنوعی با تمرکز بر شناسایی سوگیری‌های الگوریتمی و نیز تدوین شیوه‌نامه‌های شفاف برای ارزیابی اصالت پاسخ‌ها پیش از ارائه به کاربر است. علاوه بر این، طراحان سامانه‌ها ملزم هستند رابط‌های کاربری میز مرجع را به‌گونه‌ای بهینه سازی نمایند که مرز میان استدلال‌های ماشینی و محتوای مستند استنادی برای مراجعان دانشگاهی کاملاً تفکیک‌پذیر باشد.

با وجود دستاوردهای نظری و عملیاتی این الگو، چند محدودیت اساسی وجود دارد که مسیر پژوهش‌های آینده را تبیین می‌کند. نخستین محدودیت، تمرکز مدل تجربی بر ادراک متخصصان علم اطلاعات و کارشناسان هوش مصنوعی است؛ غیبت دیدگاه‌های تجربی مراجعان نهایی نظیر دانشجویان و استادان، تعمیم‌پذیری جنبه‌های تعاملی را مشروط می سازد. ازاین‌رو پیشنهاد می‌شود پژوهش‌های تکمیلی از طریق پیمایش‌های گسترده، ادراک و انتظارات اخلاقی مراجعان دانشگاهی را با این مدل بسنجند. محدودیت دوم به ساختار چارچوب مربوط می‌شود؛ درحالی‌که سطوح زیرین مدل مستقیماً بر مبنای داده‌های تجربی دلفی و دیمتل تدوین شده‌اند، سطوح فوقانی آن یعنی تعاملات سه‌گانه و پیامدهای نهایی شامل اعتماد و رضایت کاربران، محصول توسعه مفهومی مؤلفان بوده و مدل تجربی مستقیمی برای برازش آن‌ها اجرا نشده است. بنابراین، سنجش روابط علی و آزمون روابط تجربی سطوح مفهومی از طریق مدل‌سازی معادلات ساختاری[1] و همچنین تعیین وزن نهایی شاخص‌ها با استفاده از تحلیل شبکه فازی[2] پیشنهاد می‌گردد تا ابعاد عملکردی چارچوب با رویکرد تجربی جامع‌تری تثبیت شود.

---

[1]. Structural Equation Modeling (SEM)

[2]. Dematel-based Analytic Network Process (DANP)

حمایت مالی

این پژوهش هیچ‌گونه کمک مالی مشخصی از سوی هیچ‌یک از نهادهای تأمین‌کننده در بخش‌های دولتی، تجاری یا غیرانتفاعی دریافت نکرده است.

مشارکت نویسندگان

مفهوم‌سازی: [رضا کریمی] – توسعه ایده پژوهش، تدوین چارچوب اخلاقی و مفهوم‌سازی نظری مدل چهارسطحی. روش‌شناسی: [رضا کریمی] و [راحیل کریمیان] – طراحی فرایندهای دلفی فازی و دیمتل، تدوین پرسشنامه و پروتکل تحلیلی. گردآوری و آماده‌سازی داده‌ها: [راحیل کریمیان] – جستجوی پیشینه پژوهش، هماهنگی پنل خبرگان و مدیریت داده‌ها. تحلیل رسمی: [راحیل کریمیان] – انجام محاسبات دلفی فازی و دیمتل فازی در نرم‌افزار اکسل، محاسبه شاخص‌های D، R، D+R و D−R، و بررسی سازگاری. پژوهش و بررسی: [رضا کریمی] و [راحیل کریمیان] – جذب خبرگان، کدگذاری بازخوردهای کیفی و تفسیر یادداشت‌های پنل. نگارش – پیش‌نویس اولیه: [حدیثه حیدری] – تهیه نسخه اولیه دست‌نوشته و چکیده مبسوط. نگارش – بازبینی و ویرایش: [حدیثه حیدری] – بازنگری انتقادی، ویرایش زبانی و تأیید نهایی. مصورسازی: [راحیل کریمیان و حدیثه حیدری] – تهیه جداول. نظارت: [رضا کریمی] – نظارت بر پروژه و تضمین کیفیت. مدیریت پروژه: [حدیثه حیدری] – هماهنگی و زمان‌بندی.

اظهارنامه استفاده از هوش مصنوعی

نویسندگان بدین‌وسیله اعلام می‌کنند که از ابزارهای هوش مصنوعی برای تولید، تحلیل یا تفسیر داده‌های پژوهش، و همچنین برای تدوین محتوای علمی اصلی، بیان یافته‌ها یا استنتاج نتایج استفاده نشده است. ابزارهای زبانی مبتنی بر هوش مصنوعی صرفاً جهت ویرایش‌های زبانی محدود—شامل بررسی دستور زبان، بهبود خوانایی و اصلاح عبارات انگلیسی—در بخش‌هایی از متن که توسط خود نویسندگان نگارش شده بود، به کار گرفته شده‌اند. تمامی ویرایش‌های انجام‌شده با کمک هوش مصنوعی توسط نویسندگان بازبینی، تأیید و تصویب شده‌اند؛ لذا نویسندگان مسئولیت کامل صحت، یکپارچگی و اصالت کل متن پژوهش را بر عهده می‌گیرند. هیچ‌گونه ابزار هوش مصنوعی در طراحی پژوهش، فرآیند پنل خبرگان، محاسبات روش‌های «دلفی‌فازی» یا «دیمتل‌فازی»، و یا تدوین چارچوب اخلاقی دخالتی نداشته است.

تضاد منافع

هیچ تضاد منافعی وجود ندارد.

ملاحظات اخلاقی

این مطالعه مطابق با اصول اخلاقی تحقیقات شامل شرکت‌کنندگان انسانی انجام شد. همه اعضای هیئت متخصص داوطلبانه شرکت کردند و قبل از مشارکت در هر دو دور دلفی و مرحله مقایسه زوجی دیمتل‌فازی رضایت آگاهانه خود را ارائه دادند. شرکت‌کنندگان از هدف مطالعه، ماهیت داوطلبانه مشارکت خود و حق انصراف در هر مرحله بدون هیچ پیامدی مطلع شدند. در راستای تعهدات محرمانگی تحقیق، ناشناس بودن هر ۱۸ متخصص در طول جمع‌آوری داده‌ها، تجزیه و تحلیل و گزارش حفظ شد. هیچ اطلاعات هویتی در مجموعه داده‌ها ثبت یا در نسخه خطی افشا نشد. پاسخ‌های متخصصان قبل از تجزیه و تحلیل جمع‌آوری

شد تا قضاوت‌های فردی قابل ردیابی نبا شد. این مطالعه شامل جمعیت‌های آسیب‌پذیر، داده‌های شخصی حساس یا رویه‌های مداخله‌ای که نیاز به تأیید هیئت اخلاق سازمانی فراتر از پروتکل‌های استاندارد اخلاق تحقیق داشته باشند، نبود. با این وجود، این تحقیق مطابق با دستورالعمل‌های اخلاقی و استانداردهای ملی اخلاق تحقیق بود. این مطالعه شامل هیچ آزمودنی حیوانی نبود. هیچ مشوق مالی یا غیرمالی برای مشارکت ارائه نشد.

بیانیه دسترسی به داده‌ها

داده‌های پشتیبان یافته‌های این پژوهش، در صورت ارائه درخواستی موجه، از طریق نویسنده مسئول قابل دریافت هستند. مجموعه داده‌های تولید و تحلیل‌شده در این مطالعه — شامل داده‌های غربالگری به روش دلفی‌فازی، امتیازهای شاخص پس از فازی‌زدایی، ماتریس‌های مقایسه زوجی، ماتریس‌های روابط مستقیم و کل، و مقادیر محاسبه‌شده‌ی D، R، D+R و D−R — به دلیل تعهدات محرمانگی نسبت به پنل خبرگان و ضرورت حفظ گمنامی قضاوت‌های فردی شرکت‌کنندگان، به صورت عمومی منتشر نشده‌اند. داده‌های تجمیعیِ لازم برای بازتولید جداول و نمودارهای ارائه‌شده در این مقاله، با رعایت شرایطی همچون انعقاد توافق‌نامه اشتراک‌گذاری داده‌ها و کسب تأییدیه سازمانی، برای پژوهشگران واجد شرایط و به‌منظور اهداف علمی غیرتجاری قابل اشتراک‌گذاری هستند. درخواست‌ها باید به نویسنده مسئول ارسال شوند.

سپاسگزاری

نویسندگان، سپاس خود را از داوران گرانقدر که با مطالعه دقیق و ارائه نظرات ارزشمند، بر غنای این اثر افزودند، ابراز می‌دارند.

منابع

پوراجباری، منیره و سلامی، مریم (۱۴۰۴). مروری بر چالش‌های اخلاقی استفاده از هوش مصنوعی در خدمات کتابخانه‌ای. *سیزدهمین همایش ملی تحقیقات میان رشته‌ای در مدیریت و علوم انسانی*، تهران، ۱–۱۰. https://civilica.com/doc/2326068

راهداری، علیرضا و نصر، مصطفی (۱۳۹۶). چالش‌های اتاق فکر در ایران. *فصلنامه علمی – پژوهشی فرایند مدیریت و توسعه*، ۳۰ (۲)، ۲۳–۵۴. http://jmdp.ir/article-1-2727-fa.html

عظیمی، محمدحسن؛ نعمت‌الهی، زهرا و دخش، سارا (۱۴۰۱). شناسایی و طبقه‌بندی کاربردهای هوش مصنوعی در بخش‌ها و خدمات کتابخانه‌ها با استفاده از روش فراترکیب. *کتابداری و اطلاع‌رسانی*، ۲۵(۳)، ۵–۳۵. https://doi.org/10.30481/lis.2021.292701.1847

میلانی‌فر، نغمه؛ شعبانی، علی؛ فهیمی، حامد و نورزوی، علیرضا (۱۴۰۵). شناخت مؤلفه‌های چت‌بات هوش مصنوعی برای خدمات مرجع مجازی در کتابخانه‌های عمومی با استفاده از فراترکیب. *تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی*، (e735826). https://doi.org/10.61882/publlij.2026.2082554.1061

References

Adewojo, A. A., Olatunji, O. M., & Olalere, P. M. (2026). Effects of AI-driven tools on reference services and staff roles in academic libraries. *Reference Services Review,* 54(1), 15–30. https://doi.org/10.1108/RSR-02-2025-0008

Alala, A. G., Uzoaru, G. C., & Odikwa, N. H. (2024). Advantages and challenges of implementing artificial intelligence in academic library services. *Int'l Journal of Education Research and Scientific Development,* 5(1), 15. https://doi.org/10.59795/ijersd.v5i1.65

Anthonia, C., & Idiodi, E. O. (2024). The place of librarians' awareness and integration of artificial intelligence technologies for inclusive information provision. Journal of Educational Research on Children, *Parents & Teachers,* 5(1), 38–58. https://ercptjournal.org/

Asemi, A., & Asemi, A. (2018). Artificial intelligence (AI) application in library systems in Iran: A taxonomy study. *Library Philosophy and Practice (e-journal)*. 1840. https://digitalcommons.unl.edu/libphilprac/1840

Azimi, M. H., Nematolahi, Z. & Dakhesh, S. (2022). Identifying and Categorizing the Dimensions and Applications of Artificial Intelligence in Library Services Using Meta-Synthesis Method. *Library and Information Sciences*, 25(3), 5-35. doi: 10.30481/lis.2021.292701.1847 (in Persian)

Baker-Brunnbauer, J. (2021). Management perspective of ethics in artificial intelligence. *AI and Ethics,* 1, 173–181. https://doi.org/10.1007/s43681-020-00022-3

Chen, C. C., & Chang, C. C. (2025). AI-enhanced reference services in special libraries: A case study of the Hakka Literary Museum. *The Electronic Library,* 43(5), 715–732. https://doi.org/10.1108/EL-05-2025-0168

Cheng, C. H., & Lin, Y. (2002). Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation. *European journal of operational research*, 142(1), 174-186. https://doi.org/10.1016/S0377-2217(01)00280-6

Cox, C., & Tzoc, E. (2023). ChatGPT: Implications for academic libraries. *College & research libraries news*, 84(3), 99. https://doi.org/10.5860/crln.84.3.99

European Commission. (2024, January 24). *The Commission adopts its own approach on development and use of Artificial Intelligence*. https://commission.europa.eu/news-and-media/news/commission-adopts-its-own-approach-development-and-use-artificial-intelligence-2024-01-24_en

Habibi, A., Sarafrazi, A., & Izadyar, S. (2014). Delphi technique theoretical framework in qualitative research. *The international journal of engineering and science*, 3(4), 8-13. https://api.semanticscholar.org/CorpusID:32412546

Hussain, A. (2023). Use of artificial intelligence in the library services: Prospects and challenges. Library Hi Tech News, 40(2), 15–17. https://doi.org/10.1108/LHTN-11-2022-0125

Gajbhiye, C. K. (2024). Impact of artificial intelligence (AI) in library services. *International Journal For Multidisciplinary Research,* 6(3), 1-13.

International Federation of Library Associations and Institutions. (2024, April 23). *Information access: The Internet and AI*. IFLA. https://www.ifla.org/news/information-access-the-internet-and-ai/?utm_source=chatgpt.com

Lo, L. S. (2024). Transforming academic librarianship through AI reskilling: Insights from the GPT-4 exploration program. *The Journal of Academic Librarianship*, 50(3), 1-29. https://doi.org/10.1016/j.acalib.2024.102883

Manase Prabu, F., & Venkatesh, V. H. (2025). AI in library services: Balancing human interaction and automated support for mental health resources: A comparative study. *International Journal of Multidisciplinary Research and Modern Education*, 11(1), 23–29. https://doi.org/10.5281/zenodo.14675082

Mensah, G. B. (2023). Artificial intelligence and ethics: A comprehensive review of bias mitigation, transparency, and accountability in AI systems. *Preprint,* 10(1), 1–26. https://doi.org/10.13140/RG.2.2.23381.19685/1

Milanifar, N, Shabani, A, Fahimi, H and Noruzi, A. (2026). Identifying Artificial Intelligence Chatbot Components for Virtual Reference Services in Public Libraries: A Meta-Synthesis Approach. (e735826). Research on Information Science and Public Libraries, (), e735826 doi: 10.61882/publlij.2026.2082554.1061 (in Persian)

Mohanraj, A., Viji, C., Varadarajan, M. N., Kalpana, C., Jayavadivel, R., Rajkumar, N., & Jagajeevan, R. (2024). Privacy and security in digital libraries. *In AI-Assisted Library Reconstruction* (pp. 104–125). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-2782-1.ch006

OECD. (2024). *OECD AI principles*. OECD. https://www.oecd.org/en/topics/ai-principles.html?utm_source=chatgpt.com

Pourajbari, M., & Salami, M. (2025). A review of the ethical challenges of using artificial intelligence in library services. *13th National Conference on Interdisciplinary Research in Management and Humanities* (pp1-10), Tehran. https://civilica.com/doc/2326068 (in Persian)

Rahdary A., & Nasr M. (2017). Challenges of Think Tanks in Iran. *JMDP*, 30(2), 23-54 . http://jmdp.ir/article-1-2727-fa.html

Tella, A. (2023). Application of artificial intelligence for reference services in academic libraries: A global overview through a systematic review of literature. *Journal of Library Resource Sharing,* 32(1–5), 11–26. https://doi.org/10.1080/26915979.2023.2281668 (in Persian)

UNESCO. (2023). *Recommendation on the ethics of artificial intelligence*. UNESCO. https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence?hub=66973&utm_source=chatgpt.com

Yang, S. Q. (2024). ChatGPT: Unleashing the power of conversational AI for library reference services. *International Journal of Librarianship,* 9(1), 109–115. https://doi.org/10.23974/ijol.2024.vol9.1.375

Zhou, Q., Huang, W., & Zhang, Y. (2011). Identifying critical success factors in emergency management using a fuzzy DEMATEL method. *Safety Science*, 49(2), 243–252. https://doi.org/10.1016/j.ssci.2010.08.005