فرهنگ در دانشگاه اسلامی/ سال پانزدهم، شماره دوم (پیاپی۵۵)، تابستان ۱۴۰۴/صص ۸۴ – ۶۱
Culture in The Islamic University, Vol. 15, Issue.2, No.55, pp 61-84
شناسه دیجیتال: 10.22034/ciu.2025.55.61

مقاله پژوهشی اصیل
Original Article

کشف ابعاد فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی: مطالعه‌ای فراترکیب

حسین مرادی مخلص،[۱] محمدمهدی حاجیلوئی محب،[۲] امیرحسین عموئی رازانی[۳]

چکیده

هدف: هوش مصنوعی با توانایی بی‌نظیر خود در نوآوری، انقلابی در محیط‌های یادگیری ایجاد کرده است. با این حال، می‌بایست چالشهای اخلاقی آن نیز در نظر گرفته شود. هدف این مطالعه، کشف ابعاد فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی با رویکرد فراترکیب بود. **روش:** این مطالعه بر اساس هدف، تحقیق کاربردی؛ با توجه به ماهیت داده‌ها، تحقیق کیفی و با توجه به اجرای آن، جزء پژوهشهای فراترکیب با رویّهٔ هفت مرحله‌ای سندولوسکی و بارسو(۲۰۰۶) بود. کلیدواژه‌های مشخص در پایگاههای اطلاعاتی مجلات نور، مگیران، سیویلیکا، گنج، علم‌نت؛ آی‌جت، اریک، اسکوپس، ساینس دایرکت و گوگل اسکالر جستجو شد و پس از چندین مرحله غربالگری، در نهایت ۲۳ مقاله انتخاب و با کدگذاری و تحلیل مضمونی و متاسنتزی به شش مؤلفه ختم شد. **یافته‌ها:** مؤلفه‌های انصاف و برابری، حفظ حریم خصوصی و امنیت، عدم سوء استفاده، استقلال و خودمختاری، شفافیت و توضیح‌پذیری و همکاری میان رشته‌ای، در این مطالعه بررسی شد. بررسی‌ها نشان داد که ادغام ملاحظات فرااخلاقی با هوش مصنوعی، محیط‌های یادگیری اخلاقی‌تر و مؤثرتری را پرورش می‌دهد. این ترکیب منحصربه‌فرد، پیشرفتهای معناداری را در فنّاوری آموزشی امکان‌پذیر می‌کند. **نتیجه‌گیری:** نتایج نشان می‌دهد که این مؤلفه‌ها با ترکیب بینش فلسفی با اجرای عملی، مبنایی برای چارچوب آموزشی متحول‌کننده فراهم می‌کنند که مربیان و یادگیرندگان را نه تنها برای استفاده از ابزارهای پیشرفتهٔ هوش مصنوعی، بلکه برای ارزیابی انتقادی خروجی‌های آنها نیز آماده کرده و در نتیجه، نسل جدیدی از متخصصان اخلاق‌مدار و آگاه به فنّاوری را پرورش می‌دهند.

واژگان کلیدی: فرااخلاق، هوش مصنوعی، محیط یادگیری، مطالعهٔ فراترکیب.

◈ تاریخ دریافت: ۱۴۰۴/۰۱/۲۶؛ تاریخ پذیرش: ۱۴۰۴/۰۵/۱۲

[۱]. دکترای تکنولوژی آموزشی، دانشیار گروه علوم تربیتی، دانشکده علوم انسانی، دانشگاه بوعلی سینا. همدان، ایران. (نویسنده مسئول). نشانی: همدان؛ چهارباغ شهید مصطفی احمدی روشن، دانشگاه بوعلی سینا، دانشکده علوم‌انسانی. نمابر: ۰۸۱۳۳۱۱۷۸۰۴ / moradimokhles@basu.ac.ir

[۲]. دکترای معارف اسلامی، استادیار گروه معارف اسلامی، دانشکده علوم انسانی، دانشگاه بوعلی سینا. همدان، ایران. hajilouei@basu.ac.ir

[۳]. کارشناس ارشد تکنولوژی آموزشی؛ گروه علوم تربیتی، دانشکده علوم انسانی، دانشگاه بوعلی سینا. همدان، ایران. a.amooeirazani@ltr.basu.ac.ir

الف) مقدمه

هوش مصنوعی[1] شاهدی برای نبوغ انسانی است و با توانایی بی‌نظیر خود در نوآوری و انطباق، انقلابی در محیط‌های یادگیری[2] ایجاد کرده است. در حوزهٔ آموزش، قدرت دگرگون‌کنندهٔ هوش مصنوعی، منحصربه فرد است و فرصتهای بی‌سابقه‌ای را برای ارتقای تجربه یادگیری ایجاد کرده است. با استفاده از هوش مصنوعی، مربیان می‌توانند مسیرهای یادگیری شخصی‌سازی شده را متناسب با نیازهای منحصربه‌فرد هر دانش‌آموز ایجاد کنند و از طریق محتوای پویا و تطبیقی، تعامل و حفظ را تقویت کنند(بوتاریا، ۲۰۲۲). فراتر از سفارشی‌سازی، هوش مصنوعی ابزارهای بازخورد فوری ارائه می‌کند و به یادگیرندگان قدرت می‌دهد تا با استفاده از دانش خود در سناریوهای دنیای واقعی، پُلی بین نظریه و عمل بزنند(لی، ۲۰۲۳). این ادغام یکپارچه از فنّاوری و آموزش، نه تنها کیفیت یادگیری را غنی می‌کند؛ بلکه یادگیرندگان را به مهارتهایی مجهز می‌کند که در آیندهٔ مبتنی بر هوش مصنوعی، جایی که نوآوری و سازگاری در اولویت‌اند، پیشرفت کنند.

همان‌طور که از پتانسیل تحول‌آفرین هوش مصنوعی در آموزش شگفت‌زده می‌شویم، می‌بایست چالشهای اخلاقی آن را نیز در نظر بگیریم(هولمز و همکاران، ۲۰۲۲). ادغام هوش مصنوعی در محیط‌های یادگیری، سؤالات مهمی را دربارهٔ انصاف،[3] شفافیت[4] و مسئولیت‌پذیری[5] ایجاد می‌کند؛ سؤالاتی که مستلزم درک عمیق‌تر از مبانی اخلاقی‌ای است که این فنّاوری‌ها می‌بایست بر اساس آنها ساخته شده باشند(هوانگ، ۲۰۲۳؛ مطیع، ۱۳۹۹). اینجاست که فرااخلاق[6] وارد میدان می‌شود. فرااخلاق شاخه‌ای از فلسفه است که ماهیت، معنا و توجیه ادعاهای اخلاقی را بررسی می‌کند(آستروالاکیس و کاراگیانیس، ۲۰۲۴). بر خلاف اخلاق هنجاری که به ما می‌گوید باید چه کار کنیم یا اخلاق کاربردی که به مسائل اخلاقی خاصی می‌پردازد، فرااخلاق سؤالات اساسی را مطرح می‌کند: منظور ما از «درست» و «نادرست» چیست؟ آیا اصول اخلاقی، حقایق عینی‌اند یا توسط فرهنگ، زمینه یا باورهای شخصی شکل گرفته‌اند؟ چگونه قضاوتهای اخلاقی را توجیه کنیم؟ در زمینهٔ آموزش مبتنی بر هوش مصنوعی، فرااخلاق یک عینک حیاتی است برای بررسی اینکه چگونه استانداردهای جهانی، ترجیحات ذهنی یا هنجارهای ساخته‌شدهٔ اجتماعی- مانند انصاف یا حریم خصوصی- در

1. Artificial Intelligence
2. Learning Environments
3. Fairness
4. Transparency
5. Accountability
6. Meta-Ethics

سیستم‌های هوش مصنوعی رعایت می‌شود(سرلک و سعیدی، ۱۳۹۹). با استفاده از نظریه‌های فرااخلاقی، مانند نتیجه‌گرایی[1](که اعمال را بر اساس نتایج آنها ارزیابی می‌کند) یا وظیفه‌شناسی[2](که بر پایبندی به قوانین اخلاقی تأکید می‌کند)، می‌توانیم پیامدهای اخلاقی ابزارهای هوش مصنوعی را به طور انتقادی ارزیابی کنیم(سالیوان و همکاران، ۲۰۲۱). برای مثال، یک نتیجه‌گرا ممکن است مزایای یادگیری شخصی‌شده را در مقابل سوگیری‌های بالقوه در الگوریتم‌های هوش مصنوعی بسنجد؛ در حالی که یک وظیفه‌شناس ممکن است بر قوانین سختگیرانه برای حفظ حریم خصوصی داده‌ها، صرف نظر از نتایج، پافشاری کند. به این ترتیب، فرااخلاق برای توسعهٔ چارچوبهای اخلاقی قوی ضروری است و تضمین می‌کند هوش مصنوعی در آموزش، نه تنها یادگیری را ارتقا می‌دهد؛ بلکه این کار را به گونه‌ای انجام می‌دهد که منصفانه، شفاف و همسو با عمیق‌ترین ارزشهای انسانی باشد.

۱. پیشینهٔ پژوهشی

برای پرداختن به چالشهای اخلاقی پیش‌گفته، درک عمیق‌تر از مبانی اخلاقی زیربنای سیستم‌های هوش مصنوعی ضروری است. فرااخلاق، به عنوان شاخه‌ای از فلسفه که ماهیت، دامنه و توجیه قضاوتهای اخلاقی را بررسی می‌کند، چارچوبی حیاتی برای این منظور فراهم می‌کند. بر خلاف اخلاق هنجاری که نحوهٔ عمل را تجویز می‌کند، یا اخلاق کاربردی که به معضلات اخلاقی خاص می‌پردازد، فرااخلاق سؤالات اساسی را مطرح می‌کند: معنای اصطلاحات اخلاقی مانند «درست» و «نادرست» چیست؟ آیا اصول اخلاقی حقایق عینی‌اند یا توسط عوامل فرهنگی، زمینه‌ای یا ذهنی شکل می‌گیرند؟ قضاوتهای اخلاقی چگونه توجیه می‌شوند؟ در زمینهٔ محیط‌های یادگیری مبتنی بر هوش مصنوعی، نظریه‌های فرااخلاقی مانند پیامدگرایی که اعمال را بر اساس نتایج آنها ارزیابی می‌کند و وظیفه‌گرایی که بر پایبندی به قوانین اخلاقی تأکید دارد، ابزارهایی را برای ارزیابی اینکه آیا سیستم‌های هوش مصنوعی با استانداردهای اخلاقی جهانی همسو می‌باشند یا تحت تأثیر هنجارهای اجتماعی قرار دارند، ارائه می‌دهند. برای مثال، رویکرد پیامدگرا ممکن است مزایای یادگیری شخصی‌سازی شده را در برابر سوگیری‌های بالقوه در الگوریتم‌های هوش مصنوعی بسنجد؛ در حالی که دیدگاه وظیفه‌گرا ممکن است پایبندی دقیق به قوانین حریم خصوصی داده‌ها را صرف نظر از نتایج در اولویت قرار دهد. علاوه بر این، پژوهشهای فرااخلاقی دربارهٔ جهان‌شمولی ارزشهای

1. Consequentialism
2. Deontology

اخلاقی، تعامل عقل و احساس در تصمیم‌گیری اخلاقی و پایه‌گذاری قضاوتهای اخلاقی در زمینه‌های تجربی یا تکاملی می‌تواند بینشهای عمیق‌تری را در طراحی سیستم‌های هوش مصنوعی که از انصاف، حریم خصوصی و اعتماد حمایت می‌کنند، ارائه دهد.

پیش از این پژوهش، چند مطالعهٔ مروری در این رابطه انجام شده، اما موضوع به صورت کامل بررسی نشده است. پالومبو و همکاران(۲۰۲۴) با مرور مطالعات از ژانویه ۲۰۱۸ تا ژوئن ۲۰۲۳ به شناسایی و ترسیم معیارهای اخلاقی عینی مستندشده پرداخته‌اند. این مطالعه نشان می‌دهد که تنها تعداد کمی از مطالعات، معیارهای عملی و قابل عملیاتی شدن را پیشنهاد کرده‌اند؛ ضمن اینکه همه با تمرکز بر تنوع، عدم تبعیض و انصاف، عمدتاً به سایر اصول بی‌توجه بوده‌اند. همچنین نویسندگان به این نتیجه رسیدند که شکاف قابل‌توجهی بین چارچوبهای اخلاقی نظری و اجرای کمّی آنها وجود دارد. لنک(۲۰۲۲) نیز ظرفیت منحصربه‌فرد انسان برای شناخت فراسطحی را بررسی کرده است که شامل توانایی تأمل و اصلاح چارچوبهای تفسیری، اصول اخلاقی و تعاملات فنّاوری خود می‌شود. هدف نویسندگان تحلیل این است که چگونه انسانها در فراتفسیر، فرااخلاق و فرافنّاوری درگیر می‌شوند و از این طریق، درک و تعامل خود را با جهان شکل می‌دهند. این مقاله با رویکردی فلسفی و نظری به بررسی پیامدهای این فراظرفیتها بر رفتار انسان و توسعهٔ اجتماعی می‌پردازد. یافته‌ها نشان می‌دهند که این تعامل در سطح فراسطحی برای سازگاری انسان و تکامل فرهنگ و فنّاوری اساسی است. این نتیجه‌گیری بر اهمیت شناخت و پرورش این فراظرفیتها برای رسیدگی به چالشهای پیچیدهٔ اخلاقی و فنّاوری در جامعهٔ معاصر تأکید می‌کند.

پیامدهای اخلاقی هوش مصنوعی در آموزش طی سالهای اخیر توجه فزاینده‌ای را به خود جلب کرده است. فو و ونگ(۲۰۲۴) در یک بررسی نظام‌مند با تجزیه و تحلیل ۴۰ مطالعهٔ تجربی؛ انصاف، حریم خصوصی، عدم آسیب‌رسانی، استقلال و شفافیت را به عنوان ویژگی‌های حیاتی هوش مصنوعی انسان‌محور و مسئول در آموزش شناسایی کردند. ون‌نورن(۲۰۲۳) نیز در خصوص توصیهٔ یونسکو دربارهٔ اخلاق هوش مصنوعی، بر حقوق بشر، کرامت و اصولی مانند شفافیت و انصاف تأکید دارد و زمینه‌های اقدام سیاسی را برای آموزش ارائه می‌دهد. مطالعاتی مانند مطالعهٔ آیکن و اپستین(۲۰۰۰)، اصول فرااخلاقی را برای هوش مصنوعی در آموزش پیشنهاد کرده‌اند و نشان می‌دهند که هوش مصنوعی نباید دانش‌آموزان را در ابعاد اخلاقی، اجتماعی یا فکری تضعیف کند، بلکه باید قابلیتهای آنها را افزایش دهد. آثار جدیدتر، مانند کاوش فرااخلاق در اخلاق هوش مصنوعی(بودینگتون، ۲۰۲۳)،

اهمیت سؤالات فرااخلاقی را در درک قضاوتهای اخلاقی در سیستم‌های هوش مصنوعی برجسته کرده و از مثالهایی مانند آزمایش ماشین اخلاقی MIT برای نشان دادن مباحث اخلاقی عملی استفاده می‌کنند. با وجود این حجم رو به رشد از تحقیقات، هنوز در ترکیب نظام‌مند دیدگاههای فرااخلاقی خاص محیط‌های یادگیری مبتنی بر هوش مصنوعی شکاف وجود دارد؛ به ویژه در این باره که چگونه این دیدگاهها می‌توانند چارچوبهای اخلاقی را برای فنّاوری آموزشی شکل دهند.

۲. چارچوب نظری و مفهومی

فرااخلاق، دریچه‌ای انتقادی برای بررسی اصول اخلاقی بنیادی‌ای که باید راهنمای ادغام هوش مصنوعی در محیط‌های آموزشی باشند، فراهم می‌کند. بر خلاف اخلاق هنجاری که اقدامات خاصی را تجویز می‌کند، یا اخلاق کاربردی که به مسائل اخلاقی خاصی می‌پردازد، فرااخلاق سؤالات اساسی را مطرح می‌کند: اصطلاحات اخلاقی مانند «درست» و «نادرست» در زمینهٔ محیط‌های یادگیری مبتنی بر هوش مصنوعی به چه معناست؟ آیا اصول اخلاقی حقایق جهانی‌اند یا توسط عوامل فرهنگی، زمینه‌ای یا ذهنی شکل می‌گیرند؟ قضاوتهای اخلاقی چگونه توجیه می‌شوند؟ این سؤالات برای ارزیابی پیامدهای اخلاقی سیستم‌های هوش مصنوعی در آموزش، به ویژه در پرداختن به مسائل پیچیده‌ای مانند انصاف، حریم خصوصی، استقلال، شفافیت و پاسخگویی، ضروری‌اند.

در چارچوب نظری، نظریه‌های کلیدی فرااخلاقی، از جمله پیامدگرایی، وظیفه‌گرایی و اخلاق فضیلت، جایگاه محوری دارند و هر کدام دیدگاههای متمایزی دربارهٔ ارزیابی اخلاقی سیستم‌های هوش مصنوعی ارائه می‌دهند. پیامدگرایی، اخلاق، کاربردهای هوش مصنوعی را بر اساس نتایج آنها، مانند اینکه آیا یادگیری، تعامل یا رفاه دانش‌آموزان را افزایش می‌دهند، ارزیابی می‌کند. برای مثال، یک پیامدگرا ممکن است سیستم یادگیری شخصی‌سازی‌شدهٔ مبتنی بر هوش مصنوعی را با سنجش مزایای آموزشی آن در برابر خطرات بالقوه، مانند سوگیری الگوریتمی، ارزیابی کند. در مقابل، وظیفه‌گرایی بر پایبندی به قوانین و وظایف اخلاقی تأکید دارد و اصولی مانند احترام به حریم خصوصی و عدم تبعیض و انصاف را صرف نظر از نتایج، در اولویت قرار می‌دهد. رویکرد وظیفه‌گرا ممکن است ایجاب کند که سیستم‌های هوش مصنوعی برای حفظ حقوق دانش‌آموزان، به شدت از مقررات حفاظت از داده‌ها پیروی کنند. اخلاق فضیلت نیز بر ویژگی‌های اخلاقی(مانند قابل اعتماد بودن، مسئولیت‌پذیری و نیکوکاری) که سیستم‌های هوش مصنوعی باید داشته باشند،

تمرکز دارد و تضمین می‌کند که فنّاوری‌های هوش مصنوعی اعتماد را تقویت کنند و با ارزشهای آموزشی همسو باشند. علاوه بر این، پژوهشهای فرااخلاقی در نظریه‌های عدالت(مانند عدالت توزیعی و رویّه‌ای) به انصاف می‌پردازند؛ نظریه‌های مبتنی بر حقوق، زیربنای حریم خصوصی‌اند؛ نظریه‌های عاملیت اخلاقی به خودمختاری می‌پردازند و اخلاق معرفتی مربوط به دانش و توجیه، شفافیت و پاسخگویی را هدایت می‌کند. این نظریه‌ها در مجموع، چارچوبی قوی برای تحلیل مبانی اخلاقی هوش مصنوعی در آموزش ارائه می‌دهند؛ همان‌طور که توسط آیکن و اپستین(۲۰۰۰) برجسته شده است که اصول فرااخلاقی را پیشنهاد می‌کنند مبنی بر اینکه هوش مصنوعی در آموزش نباید دانش‌آموزان را در ابعاد اخلاقی، اجتماعی یا فکری تضعیف کند، بلکه باید حداقل یک بُعد را تقویت کند.

این چارچوب مفهومی با بینشهای میان‌رشته‌ای از فلسفه، آموزش، علوم کامپیوتر و اخلاق، غنی‌تر شده و ماهیت چندوجهی هوش مصنوعی را به رسمیت می‌شناسد. همان‌طور که در ادبیات ذکر شده است، فقدان چارچوبهای اخلاقی مشترک در هوش مصنوعی، نیاز به رویکردی ساختاریافته برای پرداختن به این چالشها را برجسته می‌کند(بودینگتون، ۲۰۲۳). چارچوب پیشنهادی از دستورالعملهای موجود، مانند ون‌نورن(۲۰۲۳) که بر حقوق بشر و کرامت و اصولی مانند شفافیت و انصاف در کاربردهای آموزشی هوش مصنوعی تأکید دارد، بهره می‌برد. با ادغام این دیدگاهها، این مطالعه با هدف توسعهٔ درک جامعی از مبانی فرااخلاقی که می‌تواند طراحی اخلاقی و استقرار هوش مصنوعی در آموزش را هدایت کند، انجام شده است.

در حالی که هوش مصنوعی با ارائهٔ مسیرهای یادگیری شخصی‌سازی‌شده، محتوای پویا و ابزارهای بازخورد فوری، نبوغ انسان را به نمایش می‌گذارد و محیط‌های یادگیری را متحول می‌کند؛ ادغام آن در آموزش، چالشهای اخلاقی قابل توجهی را ایجاد می‌کند. این چالشها بر تضمین عدالت در شخصی‌سازی مبتنی بر هوش مصنوعی، حفظ شفافیت در تصمیم‌گیری الگوریتمی و حفظ پاسخگویی در قبال تأثیرات سیستم‌های هوش مصنوعی بر یادگیرندگان متمرکز است. با وجود پتانسیل هوش مصنوعی برای افزایش تعامل، پیوند نظریه و عمل و تجهیز دانش‌آموزان برای آینده‌ای نوآورانه، پذیرش سریع این فنّاوری از توسعهٔ دستورالعملهای اخلاقی قوی پیشی گرفته است. با وجود مجموعهٔ گستردهٔ تحقیقات معاصر دربارهٔ به کارگیری هوش مصنوعی در محیط‌های یادگیری، پژوهشها عمدتاً بر ابعاد فنی و عملکردی متمرکزند و مبانی فلسفی- به ویژه آنهایی که مربوط به ماهیت اخلاقی هوش

ماشینی، آموزش و یادگیری‌اند- اغلب نادیده گرفته می‌شوند. این شکاف بر نیاز حیاتی به یک کاوش عمیق‌تر و میان رشته‌ای تأکید می‌کند که استقرار هوش مصنوعی در آموزش را با مفاهیم فرااخلاقی مرتبط می‌کند. چنین رویکردی، بینشهای دگرگون‌کننده‌ای ارائه می‌دهد که می‌تواند درک ما از سیستم‌های هوشمند و پیامدهای اخلاقی آنها را در زمینه‌های آموزشی تغییر دهد. برای پرداختن به این موضوع، مطالعهٔ حاضر مسیری را برای کشف و طبقه‌بندی ابعاد برنامه‌های هوش مصنوعی در محیط‌های یادگیری از دریچهٔ فرااخلاق آغاز کرده و قرنها تفکر فلسفی را برای کشف ماهیت اخلاقی پیچیدهٔ انسان و هوش مصنوعی دنبال می‌کند. هدف این مطالعه، کشف ابعاد فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی با رویکرد فراترکیب است. بر این اساس، این پژوهش به دنبال پاسخگویی به این سؤال است که: مقوله‌ها و مفاهیم فرااخلاقی که زیربنای کاربردهای هوش مصنوعی در محیط‌های آموزشی‌اند، کدام‌اند و چگونه می‌توانند به استقرار فرااخلاقی هوش مصنوعی در آموزش دست یافت؟

ب) روش پژوهش

این مطالعه بر اساس هدف آن، در دستهٔ تحقیقات کاربردی؛ با توجه به ماهیت داده‌های آن، در دستهٔ تحقیقات کیفی و با توجه به اجرای آن، در دستهٔ تحقیقات فراترکیبی طبقه‌بندی می‌شود. در این پژوهش به طور خاص از رویّهٔ هفت مرحله‌ای ساندلوسکی و باروسو(۲۰۰۶) برای پیشبرد فرایند فراترکیب استفاده شده است. رویکرد فراترکیب از روش مرور نظام‌مند که نتایج حاصل از مطالعات کیفی متعدد را برای تولید تفسیرهای مرتبهٔ بالاتر و ایجاد بینش نظری جدید ادغام می‌کند، استفاده می‌کند. ترتیب مراحل انجام شده در این مطالعه در شکل ۱ نمایش داده شده و به شرح ذیل است.

شکل ۱: روند روش فراترکیب ارائه‌شده توسط سندولوسکی و بارسو(۲۰۰۶)

۱. ساختارمند کردن سؤالات پژوهش: سؤال این پژوهش بر اساس مدل PICo ساختارمند شده است.(جدول ۱)

سؤال پژوهشی: مقوله‌ها و مفاهیم فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی کدام‌اند؟

جدول ۱: ساختارمند کردن سؤال پژوهش بر اساس مدل PICo

	مفهوم	مترادفها	سرفصلهای موضوعی آموزشی
جامعه	محیط‌های یادگیری	آموزش، تدریس	آموزش، تدریس
مداخله	هوش مصنوعی	هوش ماشینی/یادگیری عمیق	هوش
مقایسه	فرااخلاق	مبانی اخلاقی	
برآیند	مؤلفه	مقوله‌ها و مفاهیم	تأثیر

۲. ایجاد راهبرد جستجوی نظام‌مند: در مرحلهٔ دوم، طبق روش پریزما(۲۰۲۰) و بر اساس رویّهٔ ارائه‌شده توسط پیج و همکاران(۲۰۲۱)، مقالات مجلات معتبر در بازهٔ زمانی **۱۳۹۷** تا **۱۴۰۴** (۲۰۱۸–۲۰۲۵) با جستجو در پایگاههای اطلاعاتی داخلی؛ مجلات نور، مگیران، سیویلیکا، گنج و علم‌نت و پایگاههای خارجی: آی‌جت، اریک، اسکوپوس، ساینس دایرکت و گوگل اسکالر[1] و گردآوری شد. این مرور در فروردین ۱۴۰۴ (مارچ **۲۰۲۵**) و به روزرسانی بیشتر در اردیبهشت ۱۴۰۴ (آوریل **۲۰۲۵**) انجام شده است. جستجوی نظام‌مند با استفاده از مجموعه‌ای از کلمات کلیدی در حوزه‌های فرااخلاق، هوش مصنوعی و محیط‌های یادگیری انجام شد که طبق عبارت ذیل پیگیری شد:

Query: (Meta* AND Ethic*) AND (Artificial* AND Intelligen*) AND (Learn* OR Teach* OR educate*) AND (Environment*)

۳. معیارهای انتخاب یا عدم انتخاب مطالعات: ابتدا عناوین همهٔ مطالعات بازیابی‌شده بر اساس معیارهای از پیش تعیین‌شده بررسی شد. سپس چکیدهٔ مطالعاتی که غربالگری عنوان را پشت سر گذاشتند، با استفاده از معیارهای دقیق ارائه‌شده در جدول ۲، ارزیابی و در نهایت، تجزیه و تحلیلی عمیق از متون کامل مقالات باقی‌مانده انجام شد. مجموعه مقالات کنفرانسی، سرمقاله‌ها، تفسیرها، فصلهای کتاب و گزارشهای روزنامه از این بررسی حذف شدند.

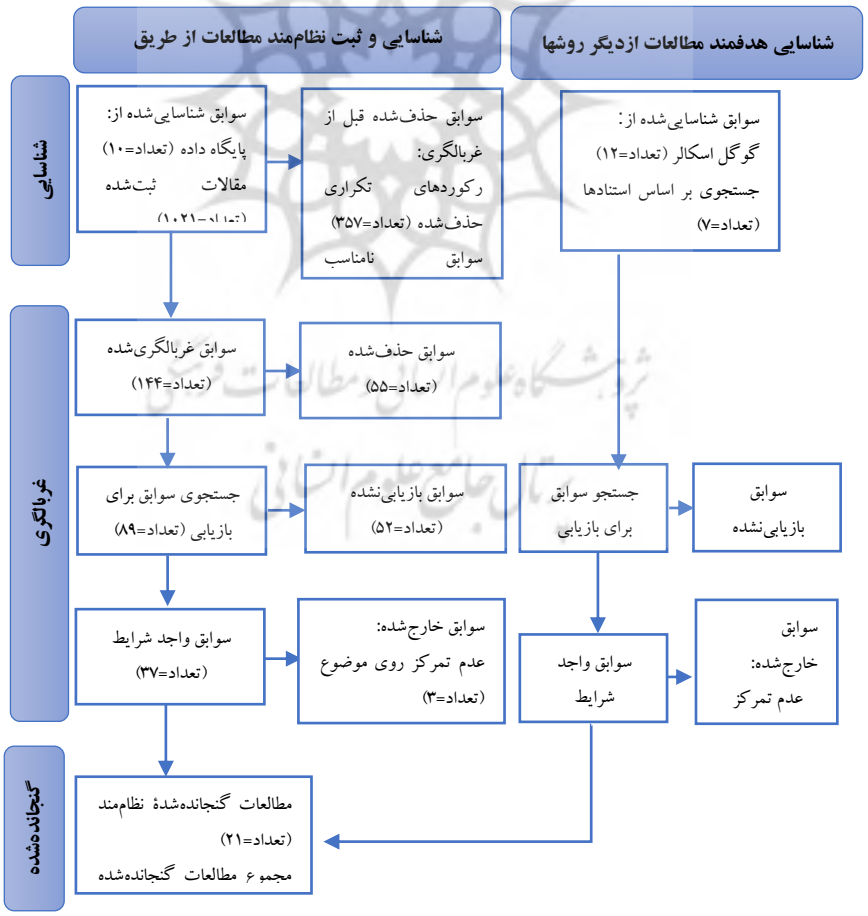
جدول ۲: معیارهای انتخاب یا عدم انتخاب مطالعات

ابعاد	معیارهای انتخاب	معیارهای عدم انتخاب

1. IJET, Eric, Scopus, Science Direct, Google Scholar

پایگاههای داده	پایگاههای اطلاعاتی داخلی: مجلات نور، مگیران، سیویلیکا، گنج و علمنت و پایگاههای خارجی: آی‌جت، اریک، اسکوپوس، ساینس دایرکت و گوگل اسکالر	منتشر شده در سایر پایگاههای دادهٔ موجود
بازه زمانی	منتشر شده در سالهای ۱۴۰۳-۱۳۹۷ شمسی	منتشر نشده در سالهای ۱۴۰۳-۱۳۹۷ شمسی
نوع مقاله	مقالات معتبر	مقالات نامعتبر

۴. فرایند انتخاب نظام‌مند مطالعات: با استفاده از کلمات کلیدی انتخاب‌شده، ابتدا ۱۰۲۱ رکورد از پایگاههای داده بازیابی شد. از این تعداد، ۲۹۱ مورد تکراری و ۵۶۷ مورد با غربالگری خودکار به عنوان نامربوط حذف شدند. پس از بررسی عنوان و چکیده، ۱۰۷ مطالعهٔ دیگر حذف شد و ۳۷ مقاله برای ارزیابی متن کامل باقی ماند که ۲۱ مورد از آنها معیارهای انتخاب را داشتند و دو مقالهٔ اضافی از طریق انتخاب هدفمند اضافه شد. فرایند انتخاب نظام‌مند مطالعات در شکل ۲ نشان داده شده است.



شکل ۲: فرایند انتخاب نظام‌مند مطالعات براساس مدل پریزما(۲۰۲۰)

۵. فرایند استخراج نظام‌مند داده‌ها: از فرمهای استاندارد برای استخراج داده‌ها استفاده شد و نتایج وارد مایکروسافت اکسل شدند.

۶. فرایند تحلیل داده‌ها: محتوای مطالعات منتخب، کدگذاری و کدگذاری‌های انجام‌شده در نرم‌افزار MAXQDA نسخه ۲۴.۹ بررسی شد. در این تحلیل از راهبردی چند مرحله‌ای در فرایند کدگذاری استفاده شد. ابتدا از کدگذاری باز برای شکستن گزارشهای مطالعه به مفاهیم مجزا و اختصاص برچسبهای اولیه استفاده شد. در مرحلهٔ بعد، کدگذاری محوری، کدهای مرتبط را با بررسی روابط و زمینه‌های آنها در دسته‌های مرتبه بالاتر دسته‌بندی کرد و در نهایت، کدگذاری انتخابی آن دسته‌ها را پالایش و ادغام کرد و مؤلفه‌های نهایی را به دست آورد.

۷. کنترل سوگیری فرایند نظام‌مند: برای اطمینان از فرایند نظام‌مند در برابر سوگیری، سه بازبین مستقل در هر مرحله درگیر شدند. بازبین نخست، استخراج داده‌ها و دسته‌بندی اولیه را مطابق با سؤالات پژوهش انجام داد؛ بازبین دوم، صحّت و کامل بودن استخراجها را تأیید کرد و همچنین عناوین، چکیده‌ها و متون کامل را با توجه به معیارهای انتخاب، مجدداً ارزیابی کرد و هرگونه مغایرت توسط بازبین سوم حل شد. نرم‌افزار Perish or Publish نسخه ۸ در تمامی مراحل توسط بازبینان استفاده شد. پایایی بین ارزیاب، اندازه‌گیری شده توسط کاپا کوهن، ۰.۶۹۴ بود که نشان دهندهٔ توافق قابل توجهی است. برای دقت بیشتر، از ابزار ریسک سوگیری کاکرین در کنار ابزار ارزیابی مقالهٔ بتن برای ارزیابی کیفیت مطالعه استفاده شد. به طور کلی، همهٔ مطالعات، روش‌شناسی صحیحی را نشان دادند؛ اگرچه تعدادی از آنها را می‌توان در انتخاب شرکت‌کنندگان، جمع‌آوری داده‌ها و روشهای تحلیلی تقویت کرد.

ج) یافته‌ها

جدول ۳، مقالات انتخاب‌شده و جدول ۴، مفاهیم و کدهای استخراج‌شدهٔ فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی را نشان می‌دهند.

جدول ۳: مقالات انتخاب‌شده

شماره	نویسنده	سال انتشار	هدف	روش	کیفیت	مفاهیم مرتبط

۱	لنک	۲۰۲۲	افق روش‌شناختی «مفسر طرح» نویسنده را از «رفتن به سطح فراسطح» به عنوان نشانه‌ای از انسانیت از طریق اخلاق بشردوستانه به سمت مشکلات مسئولیت کاربردی در فنّاوری و جامعه‌ای که از نظر فنی-انسان‌گرایانه تفسیر می‌شود، ارائه می‌کند.	کیفی	>۸۰	۱ و ۶
۲	هوانگ	۲۰۲۵	در این مقاله، فرااخلاق به عنوان یک نظریهٔ هنجاری برای عینیت ویژگی‌های اخلاقی خوبی و بدی افراد انسانی استدلال می‌کند.	کیفی	>۸۰	۱، ۲ و ۵
۳	بودینگتون	۲۰۲۳	بررسی مسائل فلسفی بیشتر زیربنای سؤالات در اخلاق هوش مصنوعی با تمرکز بر فرااخلاق، همچنین بررسی مسائل نظریهٔ دانش، فلسفه زبان و متافیزیک.	کیفی	>۸۰	۱، ۲ و ۳
۴	استنسون	۲۰۲۴	بررسی هوش مصنوعی به عنوان یک مفهوم هنجاری قدرتمند که برای اخلاق مشکل‌ساز است؛ زیرا مسائل پیچیدهٔ فلسفی را در تعریف ساده‌تر می‌کند و سطحی از دانش و اطمینان اخلاقی را فرض می‌کند که ممکن است توجیه‌پذیر نباشد.	کیفی	>۸۰	۲، ۳ و ۵
۵	کایپر	۲۰۲۱	بررسی مسائل فرااخلاق در علوم جدید و هوش مصنوعی.	کیفی	>۸۰	۱، ۲ و ۴
۶	هالاما و کالیوکوسکی	۲۰۲۲	بحث اخلاق به عنوان نقطه مرجعی برای سنجش رویکردهای فرااخلاقی و روش‌شناختی اتخاذ شده در اخلاق هوش مصنوعی.	کیفی	>۸۰	۱، ۲، ۳ و ۴
۷	وو و همکاران	۲۰۲۳	ارائهٔ رویکردی جدید برای پرداختن به موضوع فرااخلاق با	پیمایشی	>۸۰	۲، ۳، ۴ و ۵

			معرفی یک طبقه‌بندی یا چارچوب جدید هوش مصنوعی.			
۸	آیارس و نیکولز	۲۰۲۰	بررسی نقش رد جهان‌شمولی به عنوان یک ذهنیت‌گرایی در آموزش.	کیفی	>۸۰	۱، ۲، ۴، ۵ و ۶
۹	استراهونیک	۲۰۱۸	بررسی مفهوم تربیت اخلاقی و حوزهٔ نظریهٔ اخلاق و پرداختن به اهمیت نظریهٔ اخلاقی برای آموزش و تربیت اخلاقی.	کیفی	>۸۰	۲ و ۴
۱۰	مندونکا و کادیلا	۲۰۲۳	بررسی فرااخلاق در فرایندهای آموزش اخلاقی و تلاش برای غلبه بر این شکاف.	کیفی	>۸۰	۲، ۳ و ۴
۱۱	مورشر	۲۰۱۸	بررسی میزان تبعیت هوش مصنوعی به عنوان یک روش آموزشی نوین، از فرااخلاق.	کیفی	>۸۰	۱، ۴ و ۵
۱۲	بورنشتاین و هوارد	۲۰۲۱	شرح رویکردهای مختلف اخلاق هوش مصنوعی و ارائهٔ مجموعه‌ای از توصیه‌های مربوط به آموزش اخلاق هوش مصنوعی.	کیفی	>۸۰	۱، ۲، ۳ و ۴
۱۳	غازولای دورال و همکاران	۲۰۲۵	بررسی دیدگاههای نوجوانان (۱۳ تا ۱۶ سال) دربارهٔ هوش مصنوعی با تمرکز بر مباحث اخلاقی.	پیمایشی	>۸۰	۱، ۳ و ۴
۱۴	هاگندورف	۲۰۲۰	شامل اصول و توصیه‌های هنجاری با هدف مهار پتانسیلهای «اختلال‌کننده» فنّاوری‌های جدید هوش مصنوعی در آموزش.	کیفی	>۸۰	۱، ۴ و ۵
۱۵	هولمز و همکاران	۲۰۲۲	معرفی مسائلی پیرامون اخلاق هوش مصنوعی در آموزش و بررسی نظرات ۱۷ شرکت‌کننده.	پیمایشی	>۸۰	۳ و ۴
۱۶	شیف	۲۰۲۲	بررسی طرحها و انتظارات در مورد چگونگی اثر هوش مصنوعی بر بخشهای سیاست از جمله آموزش	مروری	>۸۰	۲، ۴ و ۵

			و بحث پیرامون پیامدهای اجتماعی و اخلاقی هوش مصنوعی.			
۱۷	ژوبین و همکاران	۲۰۱۹	بررسی وجود یا نبود توافق جهانی دربارهٔ هوش مصنوعی در آموزش.	کیفی	>۸۰	۲، ۳ و ۴
۱۸	ان‌جی و همکاران	۲۰۲۱	بررسی اکتشافی برای مفهوم سازی هوش مصنوعی در آموزش.	مروری	>۸۰	۳ و ۴
۱۹	ایم و سو	۲۰۲۵	بررسی سطح سواد هوش مصنوعی و کاربرد آن در آموزش.	مروری	>۸۰	۳، ۴ و ۵
۲۰	ژانگ و همکاران	۲۰۲۳	طراحی و اجرای کارگاه آموزشی توسعهٔ سواد هوش مصنوعی در آموزش.	کیفی	>۸۰	۴ و ۵
۲۱	مرادی مخلص	۱۴۰۳	ررسی مؤلفه‌های فرا اخلاقی در محیط یادگیری مجازی.	کیفی	>۸۰	۲ و ۵
۲۲	مرادی مخلص و همکاران	۱۳۹۷	بررسی رشد تربیت اخلاقی با تلفیق مؤلفه‌های تفکر انتقادی.	کیفی	>۸۰	۱، ۲، ۳ و ۴
۲۳	بلبلی قادیکلایی و پارسانیا	۱۴۰۲	شناسایی مفاهیم کلیدی اخلاق برای استفاده از هوش مصنوعی در فنّاوری‌های دیجیتال و نسبت آن با اخلاق شکوفایی.	مروری	>۸۰	۱، ۲، ۳ و ۴

جدول ۴: مفاهیم و مقوله‌های استخراج‌شده از ابعاد تأثیرگذار فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی

شماره مفهوم	مقوله	الگو	توصیف
۱	انصاف و برابری	رفتار منصفانه	اطمینان از عدم سوگیری در سیستم‌های هوش مصنوعی علیه افراد یا گروه‌ها و ارتقای فرصتها و نتایج برابر؛ که شامل پرداختن به سوگیری‌ها در الگوریتم‌های هوش مصنوعی و تقویت جامعیت در طراحی و اجرای هوش مصنوعی است.
		فرصت برابر	ارائهٔ دسترسی برابر به ابزارها و منابع آموزشی مبتنی بر هوش مصنوعی به همهٔ کاربران ، تلاش برای از بین بردن تفاوتها در مزایای سیستم هوش مصنوعی در بین جمعیتهای مختلف.

۲	حفظ حریم خصوصی و امنیت	حفاظت از داده‌ها	حفاظت از داده‌های شخصی جمع‌آوری‌شده توسط سیستم‌های هوش مصنوعی، اطمینان از رعایت مقررات حفاظت از داده‌ها و حفظ اعتماد کاربران از طریق اقدامات امنیتی قوی.
		محرمانگی	حفظ محرمانه بودن اطلاعات کاربر پردازش‌شده توسط سیستم‌های هوش مصنوعی، جلوگیری از دسترسی غیر مجاز و سوء استفاده.
۳	عدم سوء استفاده	اجتناب از ضرر	طراحی سیستم‌های هوش مصنوعی برای جلوگیری از آسیب رساندن به کاربران، اطمینان از اینکه مداخلات هوش مصنوعی اثرات نامطلوبی بر رفاه کاربران ندارد.
		سودمندی	اطمینان از اینکه سیستم‌های هوش مصنوعی به طور مثبت به تجربیات آموزشی کاربران کمک می‌کنند و نتایج یادگیری را بدون عواقب منفی ناخواسته افزایش می‌دهند.
۴	استقلال و خودمختاری	توانمندسازی کاربران	کنترل تعاملات هوش مصنوعی را در اختیار کاربران قرار داده، به آنان اجازه می‌دهد دربارهٔ نحوهٔ استفاده از ابزارهای هوش مصنوعی در فرایندهای یادگیری خود تصمیمات آگاهانه بگیرند.
		موافقت آگاهانه	اطمینان از آگاهی کامل کاربران دربارهٔ عملکردهای سیستم هوش مصنوعی و استفاده از داده‌ها، کسب رضایت آنها قبل از جمع‌آوری داده‌ها یا مداخلهٔ هوش مصنوعی.
۵	شفافیت و توضیح‌پذیری	سوگیری در داده‌های آموزشی	قابل درک کردن عملیات سیستم هوش مصنوعی برای کاربران، ارائهٔ توضیحاتی برای تصمیمات مبتنی بر هوش مصنوعی برای ایجاد اعتماد و تسهیل تعامل آگاهانهٔ کاربر.
		فرایندهای منبع باز	اطمینان از باز و در دسترس بودن فرایندهای توسعه و استقرار هوش مصنوعی به ذی‌نفعان اجازه می‌دهد رفتارها و تصمیمات سیستم هوش مصنوعی را بررسی و درک کنند.
۶	همکاری میان‌رشته‌ای	توسعهٔ مشارکتی	درگیر کردن ذی‌نفعان مختلف، از جمله مربیان، فنّاوران، اخلاق‌دانان و سیاستگذاران، در فرایند توسعهٔ هوش مصنوعی برای ترکیب دیدگاهها و تخصصهای مختلف.
		استانداردهای مشترک اخلاقی	ایجاد دستورالعملها و استانداردهای اخلاقی مشترک در بین رشته‌ها برای اطمینان از همسویی سیستم‌های هوش مصنوعی با ارزشهای گسترده اجتماعی و اهداف آموزشی.

بر اساس تحلیل متاسنتزی انجام‌شده، شش مؤلفهٔ انصاف و برابری[1](رفتار منصفانه و فرصت برابر)، حفظ حریم خصوصی و امنیت[2](حفاظت از داده‌ها و محرمانگی)، عدم سوء استفاده[3](اجتناب از ضرر و سودمندی)، استقلال و خودمختاری[4] (توانمندسازی کاربران و موافقت آگاهانه)، شفافیت و توضیح‌پذیری[5](سوگیری در داده‌های آموزشی و فرایندهای منبع باز) و همکاری میان رشته‌ای[6](توسعه مشارکتی و استانداردهای مشترک اخلاقی) استخراج شد. رفتار منصفانه در هوش مصنوعی به حصول اطمینان از اینکه تصمیمات الگوریتمی بر اساس تبعیض، تعصبات قومی و مذهبی نباشند، اشاره دارد(جیوانولا و تیریلی ، ۲۰۲۳). پالایش فعال خروجی‌ها و ترکیب الگوریتم‌های آگاه از انصاف می‌تواند نابرابری‌ها را قبل از استقرار، کشف و اصلاح کند. فرصت برابر ایجاب می‌کند که پلتفرم‌های یادگیری مبتنی بر هوش مصنوعی، به همهٔ یادگیرندگان، صرف نظر از پیشینه، دسترسی به منابع و پشتیبانی شخصی ارائه دهند(ماراس و همکاران، ۲۰۲۲). راهبردهایی مانند یادگیری تطبیقی و سیاستهای تخصیص منابع به مقابله با نابرابری‌های تاریخی و بهبود نتایج یادگیرنده کمک می‌کنند. برای مثال، عدالت ممکن است شامل هماهنگ کردن مدلهای هوش مصنوعی برای جلوگیری از سوگیری سیستمی علیه گروه‌های جمعیتی خاص باشد.

حفاظت از داده‌ها مستلزم اجرای سازوکارهایی مانند رمزگذاری، کنترلهای دسترسی و ناشناس‌سازی است تا اطمینان حاصل شود که داده‌های آموزشی که توسط سیستم‌های هوش مصنوعی استفاده می‌شوند، ایمن و سازگار باقی می‌مانند. علاوه بر این، چارچوبهای قوی حاکمیت داده، مسئولیتهای روشنی را برای نظارت بر داده‌ها در طول چرخهٔ عمر هوش مصنوعی اختصاص می‌دهند(کامنسکی، ۲۰۲۲). محرمانه بودن مستلزم آن است که اطلاعات حساس، از جمله عملکرد تحصیلی و شناسه‌های شخصی، فقط برای ذی‌نفعان مجاز قابل دسترس باشد و صرفاً برای اهداف آموزشی مشخص استفاده شود(کومار و همکاران، ۲۰۲۱). مطابق تعهدات اخلاقی و قانونی، ارائه‌دهندگان هوش مصنوعی نمی‌توانند داده‌های یادگیرنده را فراتر از شرایط توافق شده به کار برند. نمونه‌هایی مانند پیاده‌سازی احراز هویت چند عاملی یا انجام ارزیابی‌های تأثیر حریم خصوصی می‌توانند این اصول را نشان دهند.

1. Fairness & Equity
2. Privacy & Security
3. Non-maleficence
4. Agency & Autonomy
5. Transparency & Explainability
6. Interdisciplinary Collaboration

اصل «اجتناب از ضرر» الزام می‌کند که مداخلات هوش مصنوعی در محیط‌های آموزشی به دقت تأیید شود تا از اطلاعات نادرست یا آسیب‌رسانی به یادگیرندگان از طریق توصیه‌های اشتباه یا ارزیابی‌های خودکار جلوگیری شود. نظارت مستمر و سازوکارهای ایمن برای شناسایی و کاهش اثرات نامطلوب ناخواسته بر دانش‌آموزان ضروری است(آدامز و همکاران، ۲۰۲۳). اصل سودمندی، فراتر از جلوگیری از آسیب و ترویج فعالانهٔ نتایج آموزشی مثبت است؛ مانند افزایش تعامل، دسترسی و اثربخشی آموزشی از طریق ابزارهای هوش مصنوعی طراحی‌شدهٔ اخلاقی. توسعه‌دهندگان با هماهنگ کردن عملکردهای هوش مصنوعی با اهداف آموزشی، می‌توانند اطمینان حاصل کنند که فنّاوری در خدمت منافع یادگیرندگان است.(اودوناگو، ۲۰۲۳)

توانمندسازی کاربر شامل طراحی رابط‌ها و گردشهای کاری است که به یادگیرندگان، امکان کنترل بر یادگیری با کمک هوش مصنوعی را می‌دهد و آنان را قادر می‌سازد تا پیشنهادهای تولیدشده توسط هوش مصنوعی را برای حفظ نظارت انسانی، بررسی و تنظیم کنند یا نادیده بگیرند. تعبیهٔ اصول طراحی انسان‌محور تضمین می‌کند که یادگیرنده مسئولیت یادگیری خود را به دست گیرد(پرانکل، ۲۰۲۴). رضایت آگاهانه مستلزم ارتباط شفاف دربارهٔ داده‌هایی است که جمع‌آوری می‌شوند و اینکه سیستم‌های هوش مصنوعی چگونه آنها را پردازش می‌کنند. استفاده‌های آموزشی مدّ نظر، در نتیجه به یادگیرندگان اجازه می‌دهد تا دربارهٔ انتخابهای آموزشی، مشارکت داشته باشند(گوپتا و همکاران، ۲۰۲۴). تطبیق سازوکارهای رضایت با سن و شرایط یادگیرنده و بررسی مجدد رضایت در طول چرخهٔ حیات هوش مصنوعی، به استقلال یادگیرنده بیشتر احترام می‌گذارد.

سوگیری در داده‌های آموزشی زمانی پدیدار می‌شود که مجموعهٔ داده‌ها، منعکس‌کنندهٔ تعصبات تاریخی یا اجتماعی باشند که می‌تواند توسط سیستم‌های هوش مصنوعی در سراسر زبانها و فرهنگها تداوم یابد. برای شناسایی و کاهش این سوگیری‌ها، توسعه‌دهندگان باید معیارهای ارزیابی فرهنگی و مداخلات کاهش تعصب را در طول آموزش اتخاذ کنند(کیم و همکاران، ۲۰۲۰). فرایندهای باز، نیازمند افشای روشهای توسعه، منشأ داده‌ها و منطق تصمیم‌گیری‌اند تا ذی‌نفعان بتوانند در صورت نیاز، خروجی‌های هوش مصنوعی را درک و بازتولید کنند و به چالش بکشند. چارچوبهای شفافیت مانند الزامات توضیح‌پذیری خاص ذی‌نفعان، قابلیت توضیح را در تمام مراحل طراحی و استقرار سیستم هوش مصنوعی تعبیه می‌کنند.(بالاسوبرامانیم و همکاران، ۲۰۲۳)

توسعهٔ مشارکتی؛ مربیان، متخصصان اخلاق، توسعه‌دهندگان، سیاستگذاران و کاربران نهایی را گرد هم می‌آورد تا با طراحی راه حلهایی، امکان‌سنجی فنی هوش مصنوعی را با الزامات اخلاقی متعادل کنند. مشارکتهای چندبخشی؛ تبادل دانش، ابزارهای مشترک و حلقه‌های بازخورد تکرارشونده را تقویت کرده و تضمین می‌کنند که سیستم‌های هوش مصنوعی پاسخگوی نیازهای آموزشی در حال تکامل می‌باشند(تیگارد و همکاران، ۲۰۲۳). استانداردهای اخلاقی مشترک، با هماهنگ کردن دستورالعملهای بین‌المللی برای ایجاد یک پایهٔ اخلاقی یکپارچه برای هوش مصنوعی در آموزش پدید می‌آیند. این همسویی بین رشته‌ای، اجرای مداوم خط مشی و قابلیت بهبود و همکاری جهانی را در حفظ اخلاق هوش مصنوعی تسهیل می‌کند.

د) بحث و نتیجه‌گیری

این مطالعه از رویکرد فراترکیب برای کشف ابعاد فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی استفاده می‌کند. تجزیه و تحلیل متون موجود، تناقضاتی را در تحقیقات قبلی در خصوص ادغام فرااخلاق در زمینه‌های آموزشی مبتنی بر هوش مصنوعی آشکار کرد. تأکید ناکافی بر ملاحظات فرااخلاقی در توسعه و کاربرد هوش مصنوعی در محیط‌های یادگیری، مانع از ایجاد مقوله‌ها و مفاهیم مؤثری شده است که برای درک کامل پیامدهای اخلاقی هوش مصنوعی در آموزش بسیار مهم‌اند. مؤلفه‌های انصاف و برابری، حفظ حریم خصوصی و امنیت، عدم سوء استفاده، استقلال و خودمختاری، شفافیت و توضیح‌پذیری و همکاری میان‌رشته‌ای در این مطالعه بررسی شد.

در پاسخ به سؤالات تحقیق، این مطالعه نشان می‌دهد که ادغام فرااخلاق در آموزش هوش مصنوعی، مبنایی فلسفی برای درک و حل معضلات اخلاقی فراهم می‌کند و ابعاد انسانی مانند رشد اخلاقی و فکری را افزایش می‌دهد. این مطالعه بر فرااخلاق به عنوان یک مفهوم بنیادی تأکید می‌کند و استدلال می‌کند که در ادغام هوش مصنوعی در محیط‌های یادگیری، باید ماهیت و ریشه‌ها و پیامدهای مفاهیم اخلاقی در نظر گرفته شوند. از طریق رویکرد فراترکیب، ابعاد فرااخلاقی کشف‌شده، چارچوبی قوی برای درک چالشها و فرصتهای اخلاقی ذاتی در آموزش مبتنی بر هوش مصنوعی ارائه می‌دهند. این ابعاد، یک دسته‌بندی نظام‌مند ایجاد می‌کنند که درک ما را از چگونگی همسویی هوش مصنوعی با اصول فرااخلاقی عمیق‌تر می‌کند. نتایج نشان می‌دهد که ادغام ملاحظات فرااخلاقی با هوش مصنوعی در آموزش می‌تواند به عنوان یک نیروی تحول‌آفرین عمل کند و محیط‌های

یادگیری اخلاقی‌تر و مؤثرتری را پرورش دهد. این ترکیب منحصر به فرد، پیشرفتهای معناداری را در فنّاوری آموزشی امکان‌پذیر می‌کند.

پژوهشهای همسو با این مطالعه، آنهایی‌اند که به صراحت به ملاحظات فرااخلاقی در محیط‌های یادگیری مبتنی بر هوش مصنوعی می‌پردازند و اغلب از روشهای کیفی یا آمیخته استفاده می‌کنند. آیکن و اپستین(۲۰۰۰)، فرااصول هوش مصنوعی در آموزش را پیشنهاد کردند و بر تأثیر آن بر «ابعاد وجود انسان»، از جمله ویژگی‌های اخلاقی، زیبایی‌شناختی، اجتماعی، فکری، جسمی و روانی مانند شادی تأکید داشتند. فرااصول منفی آنها بیان می‌کند که فنّاوری هوش مصنوعی در آموزش، نباید این ابعاد را کاهش دهد؛ در حالی که فرااصول مثبت پیشنهاد می‌کند باید حداقل یکی از آنها را تقویت کند و چارچوبی فلسفی برای ارزیابی پیامدهای اخلاقی هوش مصنوعی ارائه دهد. اصول خاص شامل توسعهٔ سیستم‌هایی است که نقشهای جدیدی به معلمان می‌دهد و از تجلیل سیستم‌های کامپیوتری اجتناب می‌کند، که با تحقیقات فرااخلاقی دربارهٔ ماهیت تأثیر اخلاقی طنین‌انداز می‌شود. این مطالعات، مبنایی نظری برای کشف ابعاد فرااخلاقی فراهم می‌کنند و بر مبانی فلسفی به جای مسائل اخلاقی کاربردی تمرکز دارند.

پژوهشهای ناهمسو آنهایی‌اند که بر جنبه‌های فنی و کاربردی هوش مصنوعی در آموزش تمرکز دارند و اغلب نتایج تجربی و پیاده‌سازی فنی را بر ملاحظات اخلاقی، به ویژه فرااخلاق اولویت می‌دهند. زاواکی ریشتر و همکاران(۲۰۱۹)، با بررسی ۱۴۶ مقالهٔ منتشرشده از سال ۲۰۰۷ تا ۲۰۱۸، کاربردهای هوش مصنوعی از جمله: پروفایل‌سازی و پیش‌بینی، ارزیابی و سنجش، سیستم‌های تطبیقی و شخصی‌سازی و سیستم‌های تدریس هوشمند را تجزیه و تحلیل کردند. این مطالعه، روشهای فنی مانند یادگیری ماشین را به تفصیل شرح داد. با این حال، این گزارش به شکاف قابل توجهی اشاره کرد؛ تنها دو مقاله به خطرات اخلاقی پرداخته بودند و این نشان‌دهندهٔ فقدان تأمل انتقادی دربارهٔ پیامدهای آموزشی و اخلاقی، به ویژه فرااخلاق است. این مطالعات بر روشهای تجربی، مانند کتاب‌سنجی و تحلیل محتوا، و کاربردهای فنی مانند یادگیری تطبیقی و ارزیابی هوشمند تأکید دارند؛ اما به ماهیت فلسفی اخلاق نمی‌پردازند و همین امر، آنها را با تمرکز پروژه بر فرااخلاق ناهمسو می‌کند.

با وجود نتایج به دست آمده، روند این مطالعه با محدودیتهای قابل توجهی روبه‌رو بود. یکی از چالشهای اصلی، تمایل نویسندگان کلیهٔ پژوهشها به هوش مصنوعی بود و ادبیات فلسفی مرتبط با گفتمان اخلاق هوش مصنوعی به عنوان متغیّر دوم فرض می‌شد که به طور

بالقوه، ما را به مقالاتی با گرایش بیشتر به هوش مصنوعی سوق می‌داد. این تعصب ممکن است بر گنجاندن مطالعاتی تأثیر گذاشته باشد که گرچه به حوزهٔ هوش مصنوعی بسیار مرتبط می‌باشند، اما ممکن است به طور کامل نمایانگر دیدگاههای فرااخلاقی متنوع لازم برای ارائهٔ مؤلفه‌های جامع نباشند. علاوه بر این، تعداد محدود مطالعات تجربی در این حوزهٔ تخصصی، نگرانی گسترده‌تری را در خصوص فقدان تحقیقات قوی و مبتنی بر شواهد که به ابعاد فرااخلاقی هوش مصنوعی در محیط‌های یادگیری می‌پردازند، برجسته می‌کند. مطالعات آینده باید مجموعهٔ وسیع‌تری از پایگاههای داده و طیف متنوع‌تری از دیدگاههای بین رشته‌ای را در بر گیرند تا تأثیر فرضیات قبلی را به حداقل برسانند و به ارزیابی جامع‌تری از پیامدهای فرااخلاقی، روش‌شناختی و اخلاقی ادغام هوش مصنوعی در برنامه‌های درسی آموزشی دست یابند. در نتیجه، ادغام فرااخلاق با هوش مصنوعی، فرصتی بی‌نظیر برای ایجاد انقلابی در محیط‌های یادگیری ارائه می‌دهد. این مؤلفه‌ها نه تنها به شکافهای موجود در درک فلسفی و عملی کاربردهای هوش مصنوعی می‌پردازد، بلکه راه را برای ارائهٔ چارچوب آموزشی مبتنی بر اخلاق و پویا هموار می‌کند. با ادامهٔ تکامل این حوزه، ضروری است که تحقیقات آینده، این رویکردهای یکپارچه را به طور دقیق بررسی کنند و پایهٔ تجربی را برای درک بهتر پیچیدگی‌های تعاملات انسان و هوش مصنوعی در محیط‌های آموزشی گسترش دهند.

به عنوان نتیجه‌گیری، هدف مطالعهٔ حاضر ارائهٔ مؤلفه‌های ساختارمند برای کشف ابعاد فرااخلاق در محیط‌های یادگیری مبتنی بر هوش مصنوعی با استفاده از رویکرد فراترکیب بود. پس از کدگذاری تخصصی و تحلیلهای فراترکیب، مؤلفه‌های انصاف و برابری، حفظ حریم خصوصی و امنیت، عدم سوء استفاده، استقلال و خودمختاری، شفافیت و توضیح‌پذیری و همکاری میان رشته‌ای استخراج شد. به این ترتیب، مؤلفه‌های به دست آمده شامل مراحل زمینهٔ فلسفی و جهت‌گیری اخلاقی، طراحی و توسعه، اعتبارسنجی و ارزیابی انتقادی، کاربست کنترل‌شده و ادغام در برنامهٔ درسی و ارزیابی مستمر و تجدیدنظر است. این مطالعه با گنجاندن اصول فرااخلاقی تأکید می‌کند که نوآوری آموزشی صرفاً دربارهٔ پذیرش فنّاوری نیست، بلکه دربارهٔ ارتقای درک معرفتی و اخلاقی انسان است. با نگاهی به آینده، این مؤلفه‌ها تضمین می‌کنند که هوش مصنوعی در شیوه‌های آموزشی به طور مؤثرتری استفاده شود؛ ضمن اینکه ویژگی‌های اصلی حرفه‌ای‌گری آموزشی و یکپارچگی فلسفی و اخلاقی را حفظ می‌کند. این مؤلفه‌ها که برای پاسخگویی به نیازهای متنوع مربیان

و یادگیرندگان طراحی شده‌اند، تجربیات یادگیری شخصی‌سازی‌شده و حساس به زمینه و بافت را ارائه می‌دهند. در اصل، مؤلفه‌ها با ترکیب بینش فلسفی با اجرای عملی، مبنایی برای چارچوب آموزشی متحول‌کننده فراهم می‌کنند که مربیان و یادگیرندگان را نه تنها برای استفاده از ابزارهای پیشرفتهٔ هوش مصنوعی، بلکه برای ارزیابی انتقادی خروجی‌های آنها نیز آماده می‌کند و در نتیجه نسل جدیدی از متخصصان اخلاق‌مدار و آگاه به فنّاوری را پرورش می‌دهد. این مؤلفه‌ها مربیان را قادر می‌سازند تا برنامه‌های درسی مبتنی بر هوش مصنوعی را طراحی کنند که نه تنها از فنّاوری‌های پیشرفته برای ارائهٔ بهتر محتوا بهره می‌برند، بلکه تفکر انتقادی و آگاهی فلسفی و اخلاقی را در بین یادگیرندگان نیز پرورش می‌دهند. برنامه‌ریزان و تصمیم‌گیرندگان آموزشی می‌توانند از این مؤلفه‌ها برای اطلاع‌رسانی در توسعهٔ سیاستهای راهبردی استفاده کنند و اطمینان حاصل کنند که این فنّاوری نیازهای در حال تحول بخش آموزش را برآورده می‌کند. علاوه بر این، برای سیاستمداران و سیاستگذاران، این ابعاد بر اهمیت ایجاد قوانین و دستورالعملهای جامع نظارتی و اخلاقی که پذیرش مسئولانهٔ هوش مصنوعی را ترویج می‌دهند، تأکید می‌کند. بنابر این، پژوهشهای آینده باید بر بهبود الگوریتم‌ها، افزایش تعاملات انسان و ماشین و بررسی تأثیرات اجتماعی و اخلاقی این فنّاوری‌ها متمرکز شود. در نهایت، این مؤلفه‌ها نه تنها باعث کاربست فنّاوری هوش مصنوعی در آموزش می‌شود، بلکه منجر به تدوین اصول معرفتی کارامد به منظور کاربرد اخلاقی فنّاوری‌های نوین و پرورش نسل جدیدی از متخصصان اخلاق‌مدار و آگاه به فنّاوری خواهد شد.

منابع

- بلبلی قادیکلایی، سمیه و حمید پارسانیا(۱۴۰۲). «**مروری نظام‌مند بر دلالتهای اخلاقی استفاده از هوش مصنوعی در فنّاوری‌های دیجیتال و نسبت آن با اخلاق شکوفایی**». *راهبرد اجتماعی فرهنگی*، ۱۲(۳): ۷۹۸-۷۷۱.
- سرلک، علی و فاطمه سعیدی(۱۳۹۹). «**معیارهای ارزش‌شناختی اخلاق در نهج‌البلاغه با نگاه بر مکاتب وظیفه‌گرا و غایت‌گرای اخلاقی**». *مطالعات معرفتی در دانشگاه اسلامی*، ۲۴(۸۲): ۱۶۹-۱۸۸.
- مرادی مخلص، حسین(۱۴۰۳). «**فرااخلاق و محیط یادگیری مجازی: کاوش در مرزهای نوین آموزش و اخلاق**». *روندها و دستاوردها در فنّاوری یادگیری*، ۱(۲): ۱۲۰-۹۷.
- مرادی مخلص، حسین؛ جمشید حیدری و نسیبه پوطی(۱۳۹۷). «**رشد تربیت اخلاقی با تلفیق مؤلفه‌های تفکر انتقادی**». *اخلاق در علوم و فنّاوری*، ۱۳(۳): ۱۵-۸.
- مطیع، حسین(۱۳۹۹). «**آیا فنّاوری از لحاظ ارزشی جهت دارد؟**». *مطالعات معرفتی در دانشگاه اسلامی*، ۲۴(۸۳): ۴۸۰-۴۵۹.
- Adams, C.; P. Pente; G. Lemermeyer & G. Rockwell (2023). **"Ethical principles for artificial intelligence in K-12 education"**. *Computers and Education: Artificial Intelligence*, 4: 100-131.
- Aiken, R.M. & R.G. Epstein (2000). **"Ethical guidelines for AI in education: Starting a conversation"**. *International Journal of Artificial Intelligence in Education*, 11(2): 163-176.
- Ayars, A. & S. Nichols (2020). **"Rational learners and metaethics: Universalism, relativism, and evidence from consensus"**. *Mind & Language*, 35(1): 67-89.
- Balasubramaniam, N.; M. Kauppinen; A. Rannisto; K. Hiekkanen & S. Kujala (2023). **"Transparency and explainability of AI systems: From ethical guidelines to requirements"**. *Information and Software Technology*, 15(9): 107197.
- Bhutoria, A. (2022). **"Personalized education and artificial intelligence in the United States, China, and India: A systematic review using a human-in-the-loop model"**. *Computers and Education: Artificial Intelligence*, 3: 100068.
- Boddington, P. (2023). **"Philosophy for AI Ethics: Metaethics, Metaphysics, and More"**. In: *AI Ethics: Springer Nature Singapore*, 3(1): 277-317.
- Borenstein, J. & A. Howard (2021). **"Emerging challenges in AI and the need for AI ethics education"**. *AI and Ethics*, 1(1): 61-65.
- Durall Gazulla, E.; N. Hirvonen; S. Sharma; H. Hartikainen; V. Jylhä; N. Iivari; ... & A. Baizhanova (2025). **"Youth perspectives on technology ethics: analysis of teens' ethical reflections on AI in learning activities"**. *Behaviour & Information Technology*, 44(5): 888-911.

- Fu, Y. & Z. Weng (2024). **"Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices"**. *Computers and Education: Artificial Intelligence*, 7(1): 100-306.
- Giovanola, B. & S. Tiribelli (2023). **"Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms"**. *AI & society*, 38(2): 549-563.
- Gupta, N.; K. Ali; D. Jiang; T. Fink & X. Du (2024). **"Beyond autonomy: unpacking self-regulated and self-directed learning through the lens of learner agency-a scoping review"**. *BMC Medical Education*, 24(1): 1519.
- Hagendorff, T. (2020). **"The ethics of AI ethics: An evaluation of guidelines"**. *Minds and machines*, 30(1): 99-120.
- Hallamaa, J. & T. Kalliokoski (2022). **"AI ethics as applied ethics"**. *Frontiers in computer science*, 4: 776837.
- Holmes, W.; K. Porayska-Pomsta; K. Holstein; E. Sutherland; T. Baker; SB. Shum; ... & KR. Koedinger (2022). **"Ethics of AI in education: Towards a community-wide framework"**. *International Journal of Artificial Intelligence in Education*, 1: 1-23.
- Huang, Y. (2023). **"Agent-focused moral realism: Zhu Xi's virtue ethics approach to meta-ethics"**. *Australasian Philosophical Review*, 7(2): 108-133.
- Jobin, A.; M. Ienca & E. Vayena (2019). **"The global landscape of AI ethics guidelines"**. *Nature machine intelligence*, 1(9): 389-399.
- Kamenskih, A. (2022). **"The analysis of security and privacy risks in smart education environments"**. *Journal of Smart Cities and Society*, 1(1): 17-29.
- Kim, B.; J. Park & J. Suh (2020). **"Transparency and accountability in AI decision support: Explaining and visualizing convolutional neural networks for text information"**. *Decision Support Systems*, 134: 113302.
- Kiper, J. (2021). **"Toward an evolutionary-science based metaethics"**. *Religion, Brain & Behavior*, 11(1): 79-87.
- Kumar, KS.; SAH. Nair; DG. Roy; B. Rajalingam & RS. Kumar (2021). **"Security and privacy-aware artificial intrusion detection system using federated machine learning"**. *Computers & Electrical Engineering*, 96: 107440.
- Lee, AVY (2023). **"Supporting students' generation of feedback in large-scale online course with artificial intelligence-enabled evaluation"**. *Studies in Educational Evaluation*, 77, 101250.
- Lenk, H (2022). **"Humans as "meta"-beings: Meta-interpretive, meta-ethical and meta-technical"**. *New Techno Humanities*, 2(1): 47-58.
- Marras, M.; L. Boratto; G. Ramos & G. Fenu (2022). **"Equality of learning opportunity via individual fairness in personalized recommendations"**. *International Journal of Artificial Intelligence in Education*, 32(3): 636-684.

- Mendonça, D. & S. Cadilha (2023). **"A Unique Contribution to Democratic Life: Emotions, Meta-Ethics, and Meta-Cognition in P4wC"**. In: *Cultivating Reasonableness in Education: Community of Philosophical Inquiry* (57-74). Singapore: Springer Nature Singapore.
- Morscher, E. (2018). **"Meta Ethics And Moral Education"**. *Professionals' ethos and education for responsibility*, 14: 11.
- Ng, DTK.; JKL. Leung; KWS. Chu & MS. Qiao (2021). **"AI literacy: Definition, teaching, evaluation and ethical issues"**. *Proceedings of the association for information science and technology*, 58(1): 504-509.
- O'Donoghue, K (2023). **"Learning analytics within higher education: Autonomy, beneficence and non-maleficence"**. *Journal of Academic Ethics*, 21(1): 125-137.
- Page, MJ.; JE. McKenzie; PM. Bossuyt; I. Boutron; TC. Hoffmann; CD. Mulrow; ... & D. Moher (2021). **"Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement"**. *Journal of clinical epidemiology*, 134: 103-112.
- Palumbo, G.; D. Carneiro & V. Alves (2024). **"Objective metrics for ethical AI: a systematic literature review"**. *International Journal of Data Science and Analytics*, 1-21.
- Prunkl, C. (2024). **"Human autonomy at risk? An analysis of the challenges from AI"**. *Minds and Machines*, 34(3): 26.
- Sandelowski, M.; J. Barroso & CI. Voils (2006). **"Using qualitative metasummary to synthesize qualitative and quantitative descriptive findings"**. *Research in nursing & health*, 30(1): 99-111.
- Schiff, D. (2022). **"Education for AI, not AI for education: The role of education and ethics in national AI policy strategies"**. *International Journal of Artificial Intelligence in Education*, 32(3): 527-563.
- Stenson, T. (2024). **"What does it mean to be good? The normative and metaethical problem with 'AI for good"**. *AI and Ethics*, 1-10.
- Strahovnik, V. (2018). **"Ethical education and moral theory"**. *Metodički ogledi: časopis za filozofiju odgoja*, 25(2): 11-29.
- Tigard, DW.; M. Braun; S. Breuer; K. Ritt; A. Fiske; S. McLennan & A. Buyx (2023). **"Toward best practices in embedded ethics: Suggestions for interdisciplinary technology development"**. *Robotics and Autonomous Systems*, 167, 104467.
- Wu, C.; YF. Lib & P. Bouvry (2023). **"Survey of trustworthy AI: a meta decision of AI"**. *arXiv preprint arXiv:2306.00380*.
- Yim, IHY. & J. Su (2025). **"Artificial intelligence (AI) learning tools in K-12 education: A scoping review"**. *Journal of Computers in Education*, 12(1): 93-131.
- Zawacki-Richter, O.; VI. Marín; M. Bond & F. Gouverneur (2019). **"Systematic review of research on artificial intelligence applications in higher education–where are the educators?"**. *International journal of educational technology in higher education*, 16(1): 1-27.
- Zhang, H.; I. Lee; S. Ali; D. DiPaola; Y. Cheng & C. Breazeal (2023). **"Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory**

study". *International Journal of Artificial Intelligence in Education*, 33(2): 290-324.

- Bolboli Qadikolaei, S. & H. Parsania (2023). **"A systematic review of the ethical implications of using artificial intelligence in digital technologies and its relationship with the ethics of flourishing"**. *Socio-Cultural Strategy*, 12(3): 771-798. (Persian)
- Sarlak, A. & F. Saeedi (2020). **"Ethical criteria of ethics in Nahjul-Balagha with a view to the duty-oriented and teleological schools of ethics"**. *Epistemological studies in the Islamic University*, 24(82): 169-188. (Persian)
- Moradimokhles, H. (2014). **"Meta-ethics and the virtual learning environment: exploring the new frontiers of education and ethics"**. Qua*rterly Journal of Trends and Achievements in Learning Technology*, 1(2): 97-120. (Persian)
- Moradimokhles, H.; J. Heydari, & N. Poti (2018). **"The Growth of Moral Education by Integrating Critical Thinking Components"**. *Ethics in Science and Technology*, 13(3): 8-15. (Persian)
- Motie, H. (2020**). "Is Technology Value-Oriented?"**. *Epistemological Studies at the Islamic University*, 24(83): 459-480. (Prsian)