یادگیری هوشمند و تحول مدیریت

طراحی الگوی فرهنگ ریسک به‌کارگیری هوش مصنوعی در صنعت بانکداری

سعید صحت[1*]
مهدی ابراهیمی[1]
مهدی یزدان شناس[1]
زهرا خلوتی[2]

تاریخ چاپ نهایی: ۱ اردیبهشت ۱۴۰۶
تاریخ چاپ اولیه: ۳ اردیبهشت ۱۴۰۵
تاریخ پذیرش: ۱ اردیبهشت ۱۴۰۵
تاریخ بازنگری: ۲۵ فروردین ۱۴۰۵
تاریخ ارسال: ۷ بهمن ۱۴۰۴

شیوه استناددهی: صحت، سعید.، ابراهیمی، مهدی.، یزدان شناس، مهدی.، و خلوتی، زهرا. (۱۴۰۶). طراحی الگوی فرهنگ ریسک به‌کارگیری هوش مصنوعی در صنعت بانکداری. *یادگیری هوشمند و تحول مدیریت*، ۵(۱)، ۲۳-۱.

چکیده

هدف این پژوهش ارائه یک الگوی جامع، نظام‌مند و مبتنی بر شواهد برای تبیین مؤلفه‌های فرهنگ ریسک در به‌کارگیری هوش مصنوعی در صنعت بانکداری است. این پژوهش از نوع کیفی بوده و با استفاده از روش فراترکیب انجام شده است. جامعه آماری شامل مطالعات داخلی و خارجی مرتبط با فرهنگ ریسک و هوش مصنوعی در بازه زمانی ۱۹۸۸ تا ۲۰۲۵ بوده است. در مجموع ۱۴۵ مطالعه اولیه شناسایی و پس از ارزیابی کیفی با استفاده از ابزار CASP، ۷۷ منبع برای تحلیل نهایی انتخاب شد. فرآیند تحلیل بر اساس الگوی هفت‌مرحله‌ای سندلوسکی و باروسو انجام گرفت. داده‌ها از طریق کدگذاری باز، محوری و انتخابی استخراج و در قالب الگوی سیپ (بستر، درون‌داد، فرایند و برون‌داد) طبقه‌بندی شدند. پایایی کدگذاری با استفاده از ضریب کاپا (۰.۶۳۵) تأیید گردید. یافته‌ها نشان داد که فرهنگ ریسک هوش مصنوعی در بانکداری یک سازه چندبعدی و پویا است که از تعامل چهار مؤلفه اصلی شکل می‌گیرد. در مجموع ۱۸ مفهوم کلیدی و ۸۰ کد در چهار بعد بستر، درون‌داد، فرایند و برون‌داد شناسایی شد. نتایج حاکی از آن است که عوامل زمینه‌ای مانند تعهد رهبری، حاکمیت داده و انطباق مقرراتی، نقش تعیین‌کننده‌ای در شکل‌گیری فرهنگ ریسک دارند. همچنین کیفیت داده‌ها، استانداردهای فنی و منابع انسانی به‌عنوان درون‌دادهای حیاتی، بر کارآمدی سیستم‌های هوش مصنوعی اثرگذارند. در سطح فرایندی، وجود سازوکارهای نظارتی، ارزیابی مستمر و شفافیت تصمیمات الگوریتمی ضروری است. در نهایت، این تعاملات منجر به پیامدهایی نظیر افزایش اعتماد ذی‌نفعان، بهبود عملکرد سازمانی و ارتقای مزیت رقابتی می‌شود. فرهنگ ریسک هوش مصنوعی در صنعت بانکداری یک پدیده تک‌بعدی نیست، بلکه نتیجه تعامل پیچیده میان عوامل سازمانی، داده‌ای، فرایندی و پیامدی است که بدون توجه همزمان به این ابعاد، مدیریت مؤثر ریسک‌های هوش مصنوعی امکان‌پذیر نخواهد بود.

واژگان کلیدی: فرهنگ، فرهنگ ریسک، هوش مصنوعی، بانکداری، مدیریت ریسک

مشخصات نویسندگان:
۱. دانشیار، گروه مدیریت بازرگانی، دانشکده مدیریت حسابداری، دانشگاه علامه طباطبائی، تهران، ایران
۲. دانشجوی دکتری، گروه مدیریت بازرگانی، دانشگاه علامه طباطبائی، تهران، ایران

پست الکترونیکی: sehat@atu.ac.ir

© ۱۴۰۶ تمامی حقوق انتشار این مقاله متعلق به نویسنده است.

انتشار این مقاله به‌صورت دسترسی آزاد مطابق با گواهی CC BY-NC 4.0 صورت گرفته است.

Intelligent Learning and Management Transformation

A Model of Risk Culture for Artificial Intelligence Adoption in the Banking Industry

Saied Sehhat[1]*
Mahdi Ebrahimi[1]
Mahdi Yazdanshenas[1]
Zahra Khalvati[2]

Submit Date: 27 January 2026
Revise Date: 14 April 2026
Accept Date: 21 April 2026
Initial Publish: 23 April 2026
Final Publish: 21 April 2027

How to cite: Sehhat, S., Ebrahimi, M., Yazdanshenas, M., & Khalvati, Z. (2027). A Model of Risk Culture for Artificial Intelligence Adoption in the Banking Industry. *Intelligent Learning and Management Transformation, 5*(1), 1-23.

Abstract

This study aims to develop a comprehensive, systematic, and evidence-based model to explain the components of risk culture in the adoption of artificial intelligence in the banking industry. This qualitative study was conducted using a meta-synthesis approach. The research population included domestic and international studies related to risk culture and artificial intelligence published between 1988 and 2025. A total of 145 studies were initially identified, and after quality appraisal using the CASP tool, 77 sources were selected for final analysis. The synthesis process followed the seven-step method of Sandelowski and Barroso. Data were extracted through open, axial, and selective coding and categorized based on the SIPOC-like framework including context, inputs, processes, and outputs. Inter-coder reliability was confirmed using Cohen's kappa coefficient (0.635). The findings indicate that AI risk culture in banking is a multidimensional and dynamic construct formed through the interaction of four main dimensions. A total of 18 key concepts and 80 codes were identified across context, input, process, and output dimensions. Contextual factors such as leadership commitment, data governance, and regulatory compliance play a critical role in shaping risk culture. Additionally, data quality, technical standards, and human expertise serve as essential inputs influencing AI effectiveness. At the process level, continuous monitoring, algorithmic transparency, and risk evaluation mechanisms are crucial. These interactions ultimately lead to outcomes such as increased stakeholder trust, improved organizational performance, and enhanced competitive advantage. AI risk culture in the banking industry is not a single-dimensional phenomenon but the result of a complex interaction among organizational, data-related, processual, and outcome-based factors, and effective risk management cannot be achieved without addressing all these dimensions simultaneously.

Keywords: *Culture, Risk Culture, Artificial Intelligence, Banking Industry, Risk Management*

Authors' Information:
sehat@atu.ac.ir
1. Associate Professor, Department of Business Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran
2. Ph.D. Student, Department of Business Management, Allameh Tabataba'i University, Tehran, Iran

© 2027 the authors. This is an open access article under the terms of the CC BY-NC 4.0 License.

مقدمه

گسترش سریع فناوری‌های نوین، به‌ویژه هوش مصنوعی، در دهه‌های اخیر موجب تحول بنیادین در ساختارها، فرآیندها و الگوهای تصمیم‌گیری در سازمان‌ها شده است. در این میان، صنعت بانکداری به‌عنوان یکی از پیشروترین حوزه‌ها در پذیرش فناوری‌های دیجیتال، به‌شدت تحت تأثیر کاربردهای گسترده هوش مصنوعی قرار گرفته است؛ به‌گونه‌ای که از تحلیل داده‌های کلان و اعتبارسنجی مشتریان تا کشف تقلب و مدیریت ریسک، این فناوری نقش محوری ایفا می‌کند (Samoili et al., 2020; Shavandi et al., 2025). با این حال، در کنار مزایای قابل‌توجهی همچون افزایش کارایی، کاهش هزینه‌ها و بهبود کیفیت خدمات، به‌کارگیری هوش مصنوعی با مجموعه‌ای از ریسک‌های پیچیده، نوظهور و چندلایه همراه است که ماهیت آن‌ها با ریسک‌های سنتی تفاوت اساسی دارد (Curzon et al., 2021; Islam et al., 2023). این ریسک‌ها نه‌تنها ابعاد فنی مانند خطاهای الگوریتمی یا نقض حریم خصوصی را دربرمی‌گیرند، بلکه شامل پیامدهای اخلاقی، اجتماعی و نهادی نیز می‌شوند که مدیریت آن‌ها مستلزم رویکردی فراتر از چارچوب‌های کلاسیک مدیریت ریسک است (Huang et al., 2023; Morley et al., 2021). در چنین شرایطی، مفهوم «فرهنگ ریسک» به‌عنوان یکی از مؤلفه‌های کلیدی در موفقیت مدیریت ریسک سازمانی، بیش از پیش مورد توجه قرار گرفته است. فرهنگ ریسک به مجموعه‌ای از ارزش‌ها، باورها، نگرش‌ها و رفتارهای مشترک در سازمان اشاره دارد که نحوه درک، ارزیابی و پاسخ به ریسک‌ها را شکل می‌دهد (Bockius & Gatzert, 2024; Power et al., 2024). این مفهوم فراتر از ابزارها و تکنیک‌های رسمی مدیریت ریسک بوده و به‌عنوان زیرساختی نرم و بنیادین، نقش تعیین‌کننده‌ای در اثربخشی نظام‌های کنترلی و حاکمیتی ایفا می‌کند (Beasley et al., 2024; Fraser & Simkins, 2021). پژوهش‌ها نشان داده‌اند که بسیاری از شکست‌های سازمانی نه ناشی از کمبود ابزارهای فنی، بلکه نتیجه ضعف در فرهنگ ریسک و عدم همسویی رفتارهای سازمانی با اهداف مدیریت ریسک بوده است (Miller, 2022; Raab, 2022). در این راستا، ایجاد و تقویت فرهنگ ریسک به‌عنوان یک اولویت راهبردی برای سازمان‌ها، به‌ویژه در محیط‌های پیچیده و فناورانه، مطرح شده است (Esmaeilzadeh et al., 2020; Higgins, 2020).

با ورود هوش مصنوعی به صنعت بانکداری، مفهوم فرهنگ ریسک نیز دچار تحول شده و ابعاد جدیدی به آن افزوده شده است. برخلاف سیستم‌های سنتی، سیستم‌های مبتنی بر هوش مصنوعی دارای ویژگی‌هایی مانند خودیادگیری، عدم شفافیت (Black Box)، و وابستگی شدید به داده هستند که مدیریت ریسک آن‌ها را پیچیده‌تر می‌سازد (Chen et al., 2024; Johnson, 2024). این ویژگی‌ها باعث می‌شوند که تصمیمات اتخاذشده توسط الگوریتم‌ها گاه برای کاربران و حتی توسعه‌دهندگان قابل توضیح نباشد، که این امر می‌تواند به کاهش اعتماد و افزایش ریسک‌های اخلاقی و قانونی منجر شود (Huang et al., 2023; Morley et al., 2021). در نتیجه، سازمان‌ها نیازمند نوعی فرهنگ ریسک هستند که

بتواند این پیچیدگی‌ها را درک کرده و رفتارهای سازمانی را در جهت مدیریت مسئولانه و شفاف این فناوری هدایت کند (Anderson & Lee, 2024; Johnson & Patel, 2024).

در ادبیات مدیریت ریسک، چارچوب‌های مختلفی برای تبیین و توسعه فرهنگ ریسک ارائه شده است. برای مثال، استاندارد ISO ۳۱۰۰۰ بر اهمیت یکپارچگی فرهنگ ریسک با استراتژی‌های سازمانی تأکید دارد و آن را پیش‌نیاز مدیریت ریسک مؤثر می‌داند (Ivanova, 2021). همچنین، چارچوب‌های مدیریت ریسک سازمانی (ERM) بر این نکته تأکید دارند که فرهنگ ریسک باید در تمامی سطوح سازمان نهادینه شود و به‌صورت یکپارچه با فرآیندهای تصمیم‌گیری و عملکرد سازمانی ادغام گردد (Fraser & Simkins, 2021; Levy et al., 2020). با این حال، این چارچوب‌ها عمدتاً در زمینه ریسک‌های سنتی توسعه یافته‌اند و به‌طور خاص به ریسک‌های ناشی از هوش مصنوعی و ابعاد فرهنگی مرتبط با آن کمتر پرداخته‌اند.

در سال‌های اخیر، پژوهشگران تلاش کرده‌اند تا این خلأ را با توسعه مفاهیم جدیدی مانند «فرهنگ ریسک هوش مصنوعی» پر کنند. این مفهوم به‌طور خاص به نحوه درک، پذیرش و مدیریت ریسک‌های مرتبط با هوش مصنوعی در سازمان‌ها اشاره دارد و بر اهمیت عواملی مانند اخلاق داده، شفافیت الگوریتمی، پاسخگویی و یادگیری سازمانی تأکید دارد (Johnson & Patel, 2024; Streicher, 2023). مطالعات نشان می‌دهد که سازمان‌هایی که دارای فرهنگ ریسک قوی در حوزه هوش مصنوعی هستند، توانایی بیشتری در شناسایی و مدیریت ریسک‌های نوظهور داشته و عملکرد بهتری در مواجهه با عدم قطعیت‌ها از خود نشان می‌دهند (Anderson & Lee, 2024; Hancke et al., 2024). همچنین، وجود چنین فرهنگی می‌تواند به افزایش اعتماد ذی‌نفعان، بهبود انطباق با مقررات و ارتقای مزیت رقابتی سازمان منجر شود (Beasley et al., 2024; Bockius & Gatzert, 2024).

با این وجود، شواهد نشان می‌دهد که بسیاری از سازمان‌ها، به‌ویژه در صنعت بانکداری، هنوز در مراحل اولیه توسعه فرهنگ ریسک هوش مصنوعی قرار دارند. گزارش‌ها حاکی از آن است که علی‌رغم سرمایه‌گذاری‌های گسترده در فناوری‌های هوش مصنوعی، شکاف قابل‌توجهی میان بلوغ فنی و بلوغ فرهنگی در این حوزه وجود دارد (Hancke et al., 2024; Trinetix, 2025). این شکاف می‌تواند منجر به بروز ریسک‌های پیش‌بینی‌نشده، کاهش اثربخشی سیستم‌های هوش مصنوعی و حتی بحران‌های سازمانی شود. علاوه بر این، مطالعات داخلی نیز نشان‌دهنده سطح پایین یکپارچگی فرهنگ ریسک در صنعت مالی و بانکداری هستند که ضرورت توجه بیشتر به این موضوع را برجسته می‌سازد (Roshani et al., 2022; Rostami Nouroozabad et al., 2023).

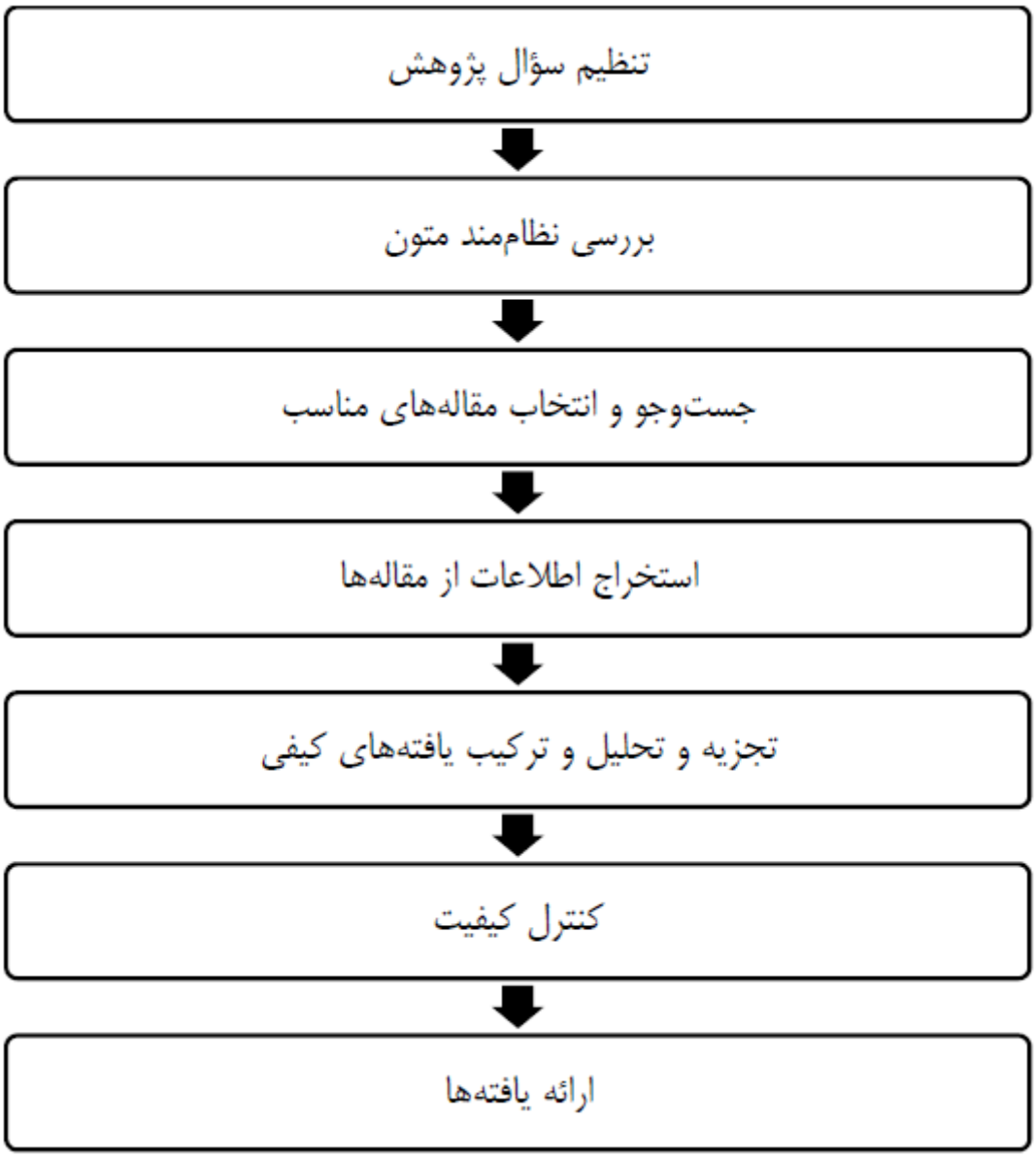
از سوی دیگر، پیچیدگی و چندبعدی بودن مفهوم فرهنگ ریسک هوش مصنوعی، نیازمند رویکردی جامع و نظام‌مند برای شناسایی و تبیین مؤلفه‌های آن است. بسیاری از مطالعات موجود، به بررسی ابعاد خاصی از این مفهوم پرداخته‌اند و کمتر پژوهشی به ارائه یک الگوی یکپارچه که

تمامی ابعاد زمینه‌ای، ورودی، فرآیندی و خروجی را به‌صورت همزمان در نظر بگیرد، اقدام کرده است (Mollaei, 2024; Power et al., 2024). همچنین، پراکندگی مفاهیم و نبود چارچوبی منسجم، موجب شده است که مدیران و سیاست‌گذاران در درک و پیاده‌سازی فرهنگ ریسک هوش مصنوعی با چالش مواجه شوند.

در این راستا، استفاده از روش‌های کیفی مانند فراترکیب می‌تواند به‌عنوان رویکردی مناسب برای تجمیع و تلفیق یافته‌های پراکنده و استخراج الگوهای مفهومی جامع مورد استفاده قرار گیرد. این روش امکان تحلیل عمیق و نظام‌مند مطالعات پیشین را فراهم کرده و به شناسایی مؤلفه‌های کلیدی و روابط میان آن‌ها کمک می‌کند. با توجه به اهمیت روزافزون هوش مصنوعی در صنعت بانکداری و نقش حیاتی فرهنگ ریسک در مدیریت مؤثر این فناوری، انجام پژوهش‌هایی که بتوانند الگوهای جامع و کاربردی در این زمینه ارائه دهند، از ضرورت بالایی برخوردار است. بر این اساس، هدف این پژوهش ارائه یک الگوی جامع و نظام‌مند برای تبیین مؤلفه‌های فرهنگ ریسک به‌کارگیری هوش مصنوعی در صنعت بانکداری بر اساس روش فراترکیب است.

روش‌شناسی

روش پژوهش حاضر کیفی است و با استفاده از روش فراترکیب به مرور نظام مند مدل‌ها و الگوهای پیشین مرتبط با فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری پرداخته است. از آنجا که مفهوم فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری در کشورهای در حال توسعه، مفهومی کمابیش جدید است و از طرفی، اغلب مقاله‌ها در این زمینه، از نوع مطالعات کیفی بوده و فاقد داده‌های کمی هستند، در پژوهش حاضر از روش فراترکیب، به عنوان روشی مناسب برای ترکیب و تلفیق جامعی از مفاهیم فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری ، بر اساس ترجمه و تفسیر مطالعات کیفی محدود استفاده شده است. به این ترتیب، پژوهش‌های گذشته (پژوهشی و مروری) در حوزهٔ فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری با استفاده از وش فراترکیب، مطابق با الگوی هفت مرحله‌ای سندلوسکی و باروسو (۲۰۰۷) بررسی شد. مراحل این بررسی در شکل ۱ خلاصه شده است.



شکل ۱. گام‌های فراترکیب منبع: سندلوسکی و باروسو (۲۰۰۷)

در این پژوهش، همان طور که در شکل ۱ نشان داده شده ، برای انجام فراترکیب از روش هفت مرحله‌ای سندلوسکی و باروسو(۲۰۰۷) استفاده شده که در ادامه هر یک از گام‌های آن تشریح شده است.

یافته‌ها

مرحله نخست: تنظیم پرسش‌های پژوهش

برای تنظیم و طرح سؤال پژوهش، از عوامل مختلفی مانند چیستی، موقعیت زمانی و چگونگی روش جامعه مورد مطالعه استفاده شد .پرسش‌های تحقیق بر اساس الگوی سیپ چنین مطرح شدند:

بسترها و زمینه‌های فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری کدام اند؟

ورودی‌ها یا درون دادهای فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری کدام اند؟

فرایند فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری چیست؟

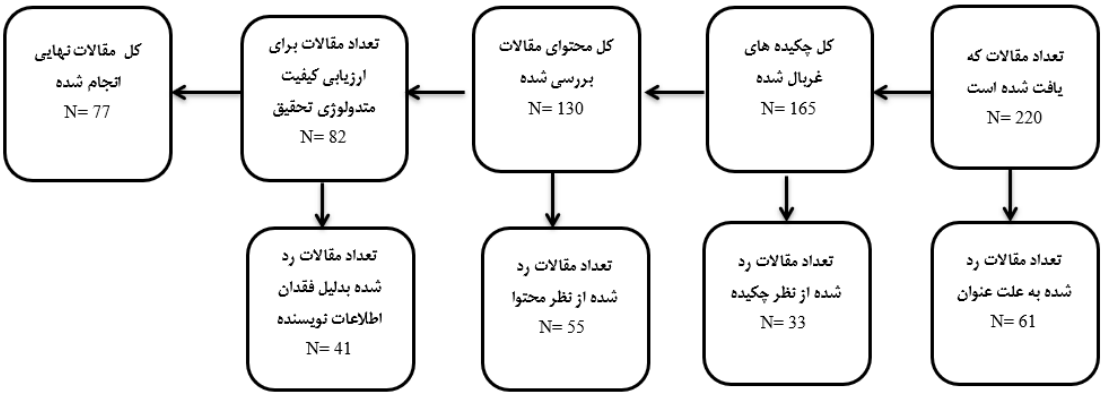
خروجی‌ها یا برون دادهای فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری کدام اند؟

مرحله دوم: بررسی نظام مند متون

منظور جست وجوی ادبیات و برای گردآوری داده‌ها از روش کتابخانه‌ای استفاده شده است.جامعهٔ آماری این پژوهش، تمام مقاله‌های منتشرشده در مجله‌های‌ای اس آی، کتاب ها، پایان نامه‌های داخلی و خارجی مرتبط با موضوع فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری در بازهٔ زمانی ۱۹۸۸ تا ۲۰۲۵ است. واژه‌های کلیدی مفهوم توسعهٔ استراتژیک فرایند ، فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری و مدل‌های مرتبط در عنوان، چکیده یا واژگان کلیدی آنها در پایگاه داده‌های داخلی نورمگز ، سیویلیکا مگ ایران ، علم نت و پایگاه داده‌های خارجی ساینس دایرکت ، امرالد ، اسکوپوس ، پروکوئست ، اسپرینگر ، وایلی، اینترساینس ، تیلور و فرانسیس و همچنین پایگاه تخصصی گوگل اسکولار جست وجو شد.

مرحله سوم: ارزیابی دقیق تحقیقات و انتخاب متون مناسب

در گام سوم، تناسب مقاله‌های یافت شده با سؤال پژوهش بررسی شد. به این ترتیب که مجموعهٔ مطالعات منتخب چندین بار بازبینی شدند. در این گام، در هر نوبت بازبینی، تعدادی از مقاله‌ها رد و از فرایند فراترکیب حذف شدند. پس از بررسی ، تناسب مقاله‌ها با عوامل مدنظر، کیفیت روش شناختی مطالعات ارزیابی شد. برنامهٔ مهارت‌های ارزیابی حیاتی یا کاسپ ابزاری برای ارزیابی کیفیت مطالعات اولیهٔ پژوهش‌های کیفی است. این ابزار از ۱۰ پرسش تشکیل میشود که مفهوم پژوهش کیفی و تعیین دقت، اعتبار و اهمیت مطالعات کیفی کمک می‌کند. درنهایت دراین مرحله ۷۷ مقاله و تعداد محدودی کتاب و پایان نامهٔ داخلی و خارجی برای تجزیه وتحلیل اطلاعات انتخاب شد. شکل ۲ خلاصه‌ای از فرایند بازبینی برای انتخاب مقاله‌های متناسب با موضوع پژوهش حاضر را نمایش می‌دهد.



شکل ۲. فرایند بازبینی و انتخاب مقاله

مرحله چهارم: استخراج اطلاعات متون

در این مرحله از فرایند فراترکیب، مقاله‌های منتخب در گام سوم، به منظور دستیابی به مؤلفه‌های فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری چندین بار مرور و مطالعه شدند. اطلاعات مربوط به هر یک از ۷۷ مقاله، شامل مؤلفه‌های توسعهٔ فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری، نویسنده و سال انتشار استخراج شد و نتایج آن در جدول درج شد.

گام پنجم: تجزیه وتحلیل و ترکیب یافته‌های کیفی

در پژوهش حاضر، ابتدا تمام مؤلفه‌های مربوط به توسعهٔ فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری شناسایی شد ، سپس بر اساس الگوی سیپ، برای هر یک از چهار مقولهٔ بستر، درون داد، مقولهٔ اصلی فرایند و برون داد، ۱۸ مفهوم و ۸۰ کد، به منظور پاسخ به پرسش‌های پژوهش استخراج شد که به ترتیب در جدول‌های ۱ تا ۴ دسته بندی شدند.

جدول ۱. مقولهٔ بستر (زمینه) الگوی فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری

مقوله	مفهوم (سنتز یافته‌ها)	کدهای مربوط (مفاهیم استخراجی)	منابع
بستر (زمینه)	بلوغ حاکمیت داده و مدل‌های AI	تعریف مالکیت داده‌ها	ISO/IEC ۲۷۰۰۱:۲۰۲۲; ECB (۲۰۲۴); Accenture (۲۰۲۳); Gholi Pour, A., et al. (۱۳۹۶)
		استانداردسازی و اعتبارسنجی مستندات مدل	NIST AI RMF; Deloitte (۲۰۲۳); Streicher, B., et al. (۲۰۱۸); FICO (۲۰۲۲)
		الگوهای نظارت بر سوگیری و عدالت	Chen, L., et al. (۲۰۲۴); Morley, J., et al. (۲۰۲۱); Huang, C., et al. (۲۰۲۳); Weston, H., et al. (۲۰۱۸)
		تفکیک نقش‌ها و مسئولیت‌ها بین تیم‌های ریسک، انطباق و توسعه‌دهندگان AI	Power, M., et al. (۲۰۲۴); FSB (۲۰۱۴); Higgins, J. (۲۰۲۰); Sinha, V. K., & Arena, M. (۲۰۲۰)
	زیرساخت فنی و یکپارچگی سامانه‌ها	سطح به‌روز بودن و مقیاس‌پذیری معماری ابری و پردازشی	Kim, et al. (۲۰۲۰); Hancke, C. H., et al. (۲۰۲۴); Johnson, T., & Patel, R. (۲۰۲۴); World Economic Forum (۲۰۲۵)

	یکپارچگی سامانه‌های مدیریت ریسک (GRC) با پلتفرم‌های AI	COSO (۲۰۱۷)؛ اسماعیل زاده، پ.، و همکاران (۱۳۹۹)؛ MIT (۲۰۲۴)؛ Islam, M., et al. (۲۰۲۳)
	کیفیت، تمیزی و قابلیت دسترس‌پذیری داده‌های آموزشی (Training Data Quality)	[IPCC (۲۰۲۱); BECG (۲۰۲۲); Trinetix (۲۰۲۵) *اشاره به مدل فرضی*]; Raab, W. L. (۲۰۲۲)
	ظرفیت سازمان برای شبیه‌سازی و تست سناریوهای ریسک	Kaplan, R. S., & Mikes, A. (۲۰۱۳); Bockius, H., & Gatzert, N. (۲۰۲۴); COSO (۲۰۲۰); Miller, P. (۲۰۲۲)
تعهد استراتژیک و تخصیص منابع	میزان سرمایه‌گذاری مالی و غیرمالی مدیریت ارشد در حوزه AI و ریسک	Agarwal, R., & Kallapur, S. (۲۰۱۸); KPMG (۲۰۰۸); Higgins, J. (۲۰۲۰); Park, Y. J. (۲۰۱۹)
	تعریف شاخص‌های کلیدی عملکرد (KPIs) مرتبط با ریسک AI در سطح هیئت مدیره	Beasley, M. S., et al. (۲۰۲۴); COSO (۲۰۲۰); Weston, H., et al. (۲۰۱۸); Miller, P. (۲۰۲۲)
	وجود فرآیندهای رسمی برای اخذ تأییدیه حاکمیتی پیش از استقرار	Power, M. (۲۰۱۶); FSB (۲۰۱۴); Hancke, C. H., et al. (۲۰۲۴); Weston, H., et al. (۲۰۱۸)
	در دسترس بودن نیروی انسانی متخصص در تقاطع AI، ریسک و انطباق	Anderson, R., & Lee, J. (۲۰۲۴); Accenture (۲۰۲۳); Mollaei, L. (۲۰۲۴); Johnson, T., & Patel, R. (۲۰۲۴)
الگوی فرهنگی پذیرش ریسک و نوآوری	هنجارهای سازمانی پیرامون پذیرش ریسک‌های غیرقابل پیش‌بینی (Tail Risk) ناشی از AI	Kaplan, R. S., & Mikes, A. (۲۰۱۳); Streicher, B., et al. (۲۰۱۸); Gholi Pour, A., et al. (۱۳۹۶); Weston, H., et al. (۲۰۱۸)
	نرخ یادگیری سازمانی از شکست‌های فناورانه	Schein, E. H. (۱۹۸۸); Anderson, R., & Lee, J. (۲۰۲۴); Raab, W. L. (۲۰۲۲); شوندی، م.، و همکاران (۱۴۰۳)
	میزان شفافیت داخلی در مورد محدودیت‌ها و خطاهای مدل‌های AI	Huang, C., et al. (۲۰۲۳); Morley, J., et al. (۲۰۲۱); FICO (۲۰۲۲); NIST AI RMF
	تأکید بر مسئولیت‌پذیری فردی در مدیریت ریسک‌های ناشی از استفاده روزمره از ابزارهای AI	COSO (۲۰۰۴); Sinha, V. K., & Arena, M. (۲۰۲۰); Park, Y. J. (۲۰۱۹); Miller, P. (۲۰۲۲)
انطباق با الزامات نظارتی و اخلاقی خارجی	درجه انطباق با مقررات ملی و بین‌المللی حاکم بر هوش مصنوعی (مانند قانون AI اتحادیه اروپا)	European Central Bank (۲۰۲۴); FSB (۲۰۱۴); Chen, L., et al. (۲۰۲۴); ISO/IEC ۴۲۰۰۱:۲۰۲۳
	نهادینه‌سازی بازبینی‌های دوره‌ای توسط ناظران خارجی	Power, M. (۲۰۱۶); FSB (۲۰۱۴); Higgins, J. (۲۰۲۰); PCAOB (۲۰۲۱) [*اشاره به مدل فرضی*]
	تعریف صریح دستورالعمل‌های اخلاقی سازمان در زمینه استفاده از AI در تصمیم‌گیری حساس	Huang, C., et al. (۲۰۲۳); Morley, J., et al. (۲۰۲۱); FICO (۲۰۲۲); Gholi Pour, A., et al. (۱۳۹۶)
	ادغام الزامات اخلاقی در فرآیند ارزیابی تأثیرات اجتماعی/اخلاقی (SEIA)	NIST AI RMF; Deloitte (۲۰۲۳); Weston, H., et al. (۲۰۱۸); Raab, W. L. (۲۰۲۲)

با توجه به جدول ۱، بستر (زمینه) پذیرش هوش مصنوعی در بانکداری بر پنج رکن اساسی استوار است **:**حاکمیت داده‌های بالغ که نیازمند تعریف مالکیت، مستندسازی دقیق و ممیزی‌های عدالت است؛ زیرساخت فنی یکپارچه که مقیاس‌پذیری، ارتباط با **GRC** و کیفیت داده‌های آموزشی را تضمین می‌کند؛ تعهد استراتژیک مدیریت ارشد که با تخصیص منابع و تعریف **KPIs** متناسب با ریسک **AI** تجلی می‌یابد؛ فرهنگ سازمانی

پذیرای ریسک نوآورانه که توانایی یادگیری سریع از شکست‌ها و شفافیت در مورد محدودیت‌های **AI** را دارد؛ و در نهایت، انطباق فعال با مقررات خارجی که مستلزم ادغام الزامات اخلاقی و نظارتی در فرآیندهای اصلی کسب‌وکار است.

جدول ۲. درون دادهای فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری

مقوله	مفهوم (سنتز یافته‌ها)	کدهای مربوط (مفاهیم استخراجی)	منابع
درون‌دادها (Inputs)	داده‌های خام و آموزشی	مجموعه‌های داده‌های تاریخی با کیفیت بالا و کم‌انحراف	Trinetix (۲۰۲۵); BECG (۲۰۲۲); Raab, W. L. (۲۰۲۲); Rostami, M., et al. (۱۴۰۲); Kim, et al. (۲۰۲۰)
		دسترسی به داده‌های برچسب‌گذاری شده معتبر از منابع داخلی و خارجی	Islam, M., et al. (۲۰۲۳); NIST AI RMF; Huang, C., et al. (۲۰۲۳); ISO/IEC ۴۲۰۰۱:۲۰۲۳; Mollaei, L. (۲۰۲۴)
		وجود فرآیندهای مستند شده برای ناشناس‌سازی و زدودن اطلاعات حساس	ISO/IEC ۲۷۰۰۱:۲۰۲۲; ECB (۲۰۲۴); Chen, L., et al. (۲۰۲۴); Power, M. (۲۰۱۶); Deloitte (۲۰۲۳)
		حجم کافی از داده‌های "حاشیه‌ای" (Edge Cases) برای تست مدل در شرایط استرس	Bockius, H., & Gatzert, N. (۲۰۲۴); Weston, H., et al. (۲۰۱۸); Miller, P. (۲۰۲۲); Kaplan, R. S., & Mikes, A. (۲۰۱۳); FICO (۲۰۲۲)
		مکانیزم‌های تأیید حقوق مالکیت داده‌ها قبل از استفاده در آموزش مدل	Power, M. (۲۰۱۶); FSB (۲۰۱۴); Higgins, J. (۲۰۲۰); Gholi Pour, A., et al. (۱۳۹۶); ISO/IEC ۲۷۰۰۱:۲۰۲۲
	الگوها و استانداردهای فنی	دسترسی به مدل‌های پایه از پیش آموزش‌دیده در صورت نیاز	MIT (۲۰۲۴); Accenture (۲۰۲۳); Johnson, T., & Patel, R. (۲۰۲۴); World Economic Forum (۲۰۲۵); Kim, et al. (۲۰۲۰)
		وجود کتابخانه‌های نرم‌افزاری استاندارد و تأیید شده برای توسعه و استقرار مدل (	Kim, et al. (۲۰۲۰); Islam, M., et al. (۲۰۲۳); Johnson, T., & Patel, R. (۲۰۲۴); MIT (۲۰۲۴); Trinetix (۲۰۲۵)
		دسترسی به گزارش‌های ارزیابی ریسک الگوریتمی (AIA Reports) مدل‌های قبلی یا همتا	Deloitte (۲۰۲۳); Weston, H., et al. (۲۰۱۸); Chen, L., et al. (۲۰۲۴); NIST AI RMF; ECB (۲۰۲۴)
		ابزارهای نرم‌افزاری برای محاسبه خودکار معیارهای ریسک (مثل انحراف یا مدل سنسوریتی)	FICO (۲۰۲۲); Johnson, R. (۲۰۲۴); Mollaei, L. (۲۰۲۴); Islam, M., et al. (۲۰۲۳); Anderson, R., & Lee, J. (۲۰۲۴)
		استانداردهای اعتبارسنجی مدل تعریف شده توسط واحد ریسک	COSO (۲۰۲۰); Beasley, M. S., et al. (۲۰۲۴); Hancke, C. H., et al. (۲۰۲۴); Miller, P. (۲۰۲۲); Power, M. (۲۰۱۶)
	منابع انسانی و دانش تخصصی	در دسترس بودن متخصصین داده و مهندسین یادگیری ماشین با تجربه کافی	Anderson, R., & Lee, J. (۲۰۲۴); Accenture (۲۰۲۳); Mollaei, L. (۲۰۲۴); Johnson, T., & Patel, R. (۲۰۲۴); Park, Y. J. (۲۰۱۹)

	وجود تیم‌های متخصص در حوزه انطباق (Compliance) و اخلاق هوش مصنوعی	Morley, J., et al. (۲۰۲۱); Huang, C., et al. (۲۰۲۳); Higgins, J. (۲۰۲۰); Power, M. (۲۰۱۶); شوندی، م.، و همکاران(۱۴۰۳)
	دسترسی به مشاوره حقوقی داخلی/خارجی برای تفسیر مقررات متغیر AI	ECB (۲۰۲۴); FSB (۲۰۱۴); Higgins, J. (۲۰۲۰); Beasley, M. S., et al. (۲۰۲۴); Power, M. (۲۰۱۶)
	منابع آموزشی به‌روز برای تمام کارکنان درگیر در فرایند توسعه و بهره‌برداری	Schein, E. H. (۱۹۸۸); Anderson, R., & Lee, J. (۲۰۱۹); Park, Y. J. (۲۰۲۴); شوندی، م.، و همکاران (۱۴۰۳); Miller, P. (۲۰۲۲)
	نقش‌های تعریف شده برای "تأیید کننده نهایی ریسک" توسط مدیریت ارشد (Risk Approvers)	COSO (۲۰۲۰); Kaplan, R. S., & Mikes, A. (۲۰۱۶); Beasley, M. S., et al. (۲۰۲۴); Sinha, V. K., & Arena, M. (۲۰۲۰); Power, M. (۲۰۱۶)
دستورالعمل‌های سیاست‌گذاری داخلی	سیاست‌های رسمی سازمان در مورد آستانه‌های پذیرش ریسک مرتبط با AI	COSO (۲۰۲۰); Bockius, H., & Gatzert, N. (۲۰۲۴); Kaplan, R. S., & Mikes, A. (۲۰۱۶); Rostami, M., et al. (۱۴۰۲); Beasley, M. S., et al. (۲۰۲۴)
	منشور اخلاقی هوش مصنوعی (AI Ethics Charter) مصوب هیئت مدیره	Huang, C., et al. (۲۰۲۳); Morley, J., et al. (۲۰۲۱); ISO/IEC ۴۲۰۰۱:۲۰۲۳; Gholi Pour, A., et al. (۱۳۹۶); Weston, H., et al. (۲۰۱۸)
	دستورالعمل‌های داخلی برای مستندسازی و تصمیم‌گیری‌های اتخاذ شده در چرخه حیات مدل (Model Governance Documentation)	Power, M. (۲۰۱۶); ISO/IEC ۲۷۰۰۱:۲۰۲۲; Higgins, J. (۲۰۲۰); NIST AI RMF; Power, M., et al. (۲۰۲۴)
	تعریف رسمی تعاریف کلیدی (مانند سوگیری، انصاف، شفافیت) در سطح سازمان	Chen, L., et al. (۲۰۲۴); FICO (۲۰۲۲); Miller, P. (۲۰۲۲); Weston, H., et al. (۲۰۱۸); COSO (۲۰۰۴)
	بودجه تخصیص یافته و مصوب برای پروژه‌های مدیریت ریسک و انطباق AI	KPMG (۲۰۰۸); Beasley, M. S., et al. (۲۰۲۴); Higgins, J. (۲۰۲۰); Park, Y. J. (۲۰۱۹); Hancke, C. H., et al. (۲۰۲۴)

درون‌دادهای اساسی برای الگوی فرهنگی ریسک در بانکداری مبتنی بر هوش مصنوعی، به چهار دسته حیاتی تقسیم می‌شوند: نخست، داده‌های خام و آموزشی که کیفیت، برچسب‌گذاری صحیح، حذف اطلاعات حساس **(PII)** و پوشش موارد حاشیه‌ای **(Edge Cases)** آن‌ها، مبنای سلامت الگو است؛ دوم، الگوها و استانداردهای فنی که شامل دسترسی به مدل‌های پایه استانداردشده، کتابخانه‌های نرم‌افزاری تأیید شده، و ابزارهای خودکار محاسبه‌گر ریسک بر اساس معیارهای اعتبارسنجی تعریف‌شده توسط ریسک است؛ سوم، منابع انسانی و دانش تخصصی شامل ترکیبی از مهندسین **ML** ، افسران اخلاق **AI** ، مشاوره حقوقی برای تفسیر مقررات، و نقش‌های کاملاً تعریف شده برای تأییدکنندگان نهایی ریسک؛ و در نهایت، دستورالعمل‌های سیاست‌گذاری داخلی که باید شامل منشور اخلاقی مصوب هیئت مدیره، تعریف رسمی آستانه‌های پذیرش ریسک **(RAS)**مرتبط با **AI** و سیاست‌های بودجه‌ای مشخص برای پروژه‌های انطباق باشند.

جدول ۳. فرایند فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری

مقوله	مفهوم (سنتز یافته‌ها)	کدهای مربوط (مفاهیم استخراجی)	منابع
فرایندها (مدیریت و عملیات)	فرآیند توسعه مدل مسئولانه	اجرای فرایند تفکر/طراحی اولاً اخلاقی در مرحله مفهومی‌سازی	Huang, C., et al. (۲۰۲۳); Morley, J., et al. (۲۰۲۱); ISO/IEC ۴۲۰۰۱:۲۰۲۳; NIST AI RMF
		ادغام مراحل ارزیابی ریسک‌های پنهان در حلقه توسعه (SDLC)	Islam, M., et al. (۲۰۲۳); Streicher, B., et al. (۲۰۱۸); Power, M., et al. (۲۰۲۴); Chen, L., et al. (۲۰۲۴)
		استفاده از فرآیندهای یادگیری ماشین عملیاتی (MLOps) خودکار برای به‌روزرسانی و بازآزمایی مدل‌ها	Kim, et al. (۲۰۲۰); Johnson, T., & Patel, R. (۲۰۲۴); World Economic Forum (۲۰۲۵); Accenture (۲۰۲۳)
		اجرای ارزیابی تأثیر الگوریتمی (AIA) قبل از استقرار در محیط عملیاتی	Deloitte (۲۰۲۳); ECB (۲۰۲۴); Weston, H., et al. (۲۰۱۸); FSB (۲۰۱۴)
	فرایندهای مدیریت ریسک عملیاتی	مکانیزم‌های نظارت مستمر بر انحراف مدل از عملکرد مورد انتظار	Islam, M., et al. (۲۰۲۳); Power, M. (۲۰۱۶); FICO (۲۰۲۲); Johnson, R. (۲۰۲۴)
		اجرای تست‌های حساسیت و فرضی (Scenario Testing) بر مدل‌ها تحت شرایط بحرانی غیرمنتظره	Kaplan, R. S., & Mikes, A. (۲۰۱۳); Bockius, H., & Gatzert, N. (۲۰۲۴); COSO (۲۰۲۰); Miller, P. (۲۰۲۲)
		فرآیندهای بازنگری و اعتبارسنجی مدل‌های یادگیری تقویتی	Chen, L., et al. (۲۰۲۴); Morley, J., et al. (۲۰۲۱); NIST AI RMF; ISO/IEC ۴۲۰۰۱:۲۰۲۳
		ادغام ریسک‌های سیستمی ناشی از وابستگی متقابل AI در فرایندهای جامع مدیریت ریسک بانک	COSO (۲۰۱۷); FSB (۲۰۱۴); Hancke, C. H., et al. (۲۰۲۴); Sinha, V. K., & Arena, M. (۲۰۲۰)
	فرآیندهای تعامل با نهادهای نظارتی و گزارش‌دهی	مکانیزم‌های خودکار برای تولید گزارش‌های انطباق به نهادهای نظارتی	ECB (۲۰۲۴); ISO/IEC ۲۷۰۰۱:۲۰۲۲; Power, M. (۲۰۱۶); Higgins, J. (۲۰۲۰)
		فرایندهای مشخص برای ارتباط با مشتریان در خصوص تصمیمات اتخاذ شده توسط AI (شفافیت برون‌سازمانی)	Huang, C., et al. (۲۰۲۳); Weston, H., et al. (۲۰۱۸); Raab, W. L. (۲۰۲۲); Sinha, V. K., & Arena, M. (۲۰۲۰)
		فرآیند داخلی به‌روزرسانی مدل‌ها در پاسخ مستقیم به تغییرات محیطی نظارتی	Power, M. (۲۰۱۶); Deloitte (۲۰۲۳); Higgins, J. (۲۰۲۰); ISO/IEC ۴۲۰۰۱:۲۰۲۳
		برگزاری جلسات منظم بین واحد ریسک و واحد توسعه برای همسوسازی تعاریف ریسک	COSO (۲۰۲۰); Beasley, M. S., et al. (۲۰۲۴); Kaplan, R. S., & Mikes, A. (۲۰۱۶); Miller, P. (۲۰۲۲)
	فرایندهای مدیریت تغییر و آموزش	تدوین استانداردهای عملیاتی جدید (SOPs) برای مشاغل متأثر از جایگزینی وظایف توسط AI	Schein, E. H. (۱۹۸۸); Anderson, R., & Lee, J. (۲۰۲۴); Johnson, R. (۲۰۲۴); Gholi Pour, A., et al. (۱۳۹۶)
		پیاده‌سازی برنامه‌های آموزشی اجباری برای افزایش سواد AI–ریسک در سطح مدیران میانی	Anderson, R., & Lee, J. (۲۰۲۴); Park, Y. J. (۲۰۱۹)؛شوندی، م.، و همکاران (۱۴۰۳); Mollaei, L. (۲۰۲۴)
		فرایندهای ارزیابی تأثیر شغلی ناشی از اتوماسیون AI	Power, M. (۲۰۱۶); Beasley, M. S., et al. (۲۰۲۴); Miller, P. (۲۰۲۲); Viscelli, T. R., et al. (۲۰۱۷)

	ایجاد کانال‌های بازخورد مستمر از کاربران نهایی (Front-line Users) درباره نقص‌های مدل	Weston, H., et al. (۲۰۱۸); Morley, J., et al. (۲۰۲۱); Raab, W. L. (۲۰۲۲); Anderson, R., & Lee, J. (۲۰۲۴)
فرایند تأمین داده و پیش‌پردازش	رویه‌های حاکمیتی برای تأیید منبع و کیفیت داده‌های ورودی به مدل‌ها	ISO/IEC ۲۷۰۰۱:۲۰۲۲; Trinetix (۲۰۲۵); BECG (۲۰۲۲); Raab, W. L. (۲۰۲۲)
	اجرای فرایندهای برچسب‌گذاری و نرمال‌سازی داده‌ها با نظارت انسانی	Islam, M., et al. (۲۰۲۳); NIST AI RMF; Huang, C., et al. (۲۰۲۳); Rostami, M., et al. (۱۴۰۲)
	مکانیسم‌های مستندسازی و حذف داده‌های شخصی یا تبعیض‌آمیز	ISO/IEC ۴۲۰۰۱:۲۰۲۳; Chen, L., et al. (۲۰۲۴); ECB (۲۰۲۴); FICO (۲۰۲۲)
	تعریف سقف‌های کیفیت داده که در صورت عدم دستیابی، فرایند توسعه متوقف می‌شود	BECG (۲۰۲۲); COSO (۲۰۱۷); Johnson, T., & Patel, R. (۲۰۲۴); Kim, et al. (۲۰۲۰)

با توجه به جدول ۳، فرایندها (مدیریت و عملیات) شکل‌دهنده الگوی عملیاتی برای پیاده‌سازی فرهنگ ریسک در به‌کارگیری هوش مصنوعی هستند و این امر از طریق پنج دسته فرآیندی مجزا محقق می‌شود: نخست، توسعه الگوی مسئولانه که نیازمند ادغام اصول اخلاقی از مرحله طراحی[1]، ارزیابی ریسک‌های پنهان در چرخه توسعه نرم‌افزار(**SDLC**) ، و استفاده از چارچوب‌های **MLOps** برای نظارت و بازآزمایی مستمر است؛ دوم، مدیریت ریسک عملیاتی که بر نظارت لحظه‌ای بر انحراف مدل (**Drift**) و اجرای تست‌های حساسیت سخت‌گیرانه تحت شرایط بحرانی تمرکز دارد؛ سوم، تعامل با نهادهای نظارتی که مستلزم خودکارسازی گزارش‌دهی انطباق و ایجاد شفافیت در مورد تصمیمات **AI** برای مشتریان است؛ چهارم، مدیریت تغییر و آموزش که شامل تدوین **SOP** های جدید، آموزش اجباری سواد-**AI** ریسک، و ایجاد کانال‌های بازخورد مستقیم از کاربران خط مقدم است؛ و پنجم، تأمین داده و پیش‌پردازش که نظارت انسانی بر برچسب‌گذاری[2] و تعیین سقف‌های کیفیت داده را به عنوان دروازه ورود مدل‌ها تعریف می‌کند.

جدول ۴. برون دادهای فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری

مقوله	مفهوم (سنتز یافته‌ها)	کدهای مربوط (مفاهیم استخراجی)	منابع
نتایج (برون‌ده‌ها)	اثربخشی و قابلیت اطمینان مدل‌های AI	کاهش نرخ خطا و نویز در پیش‌بینی‌های مدل‌های AI پس از اعمال حاکمیت بستر	Anderson, R., & Lee, J. (۲۰۲۴); Islam, M., et al. (۲۰۲۳); Johnson, R. (۲۰۲۴); Mollaei, L. (۲۰۲۴); MIT (۲۰۲۴)
		افزایش شفافیت و قابلیت توضیح‌پذیری (Explainability) خروجی‌های مدل برای کاربران و ناظران	Huang, C., et al. (۲۰۲۳); Morley, J., et al. (۲۰۲۱); FICO (۲۰۲۲); Chen, L., et al. (۲۰۲۴); NIST AI RMF
		کاهش زمان شناسایی و اصلاح مشکلات مدل (Time-to-Remediate)	Islam, M., et al. (۲۰۲۳); World Economic Forum (۲۰۲۵); Accenture (۲۰۲۳); Johnson, T., & Patel, R. (۲۰۲۴); Statista (۲۰۲۳)

[1] Ethical by Design
[2] Human-in-the-Loop

	بهبود در انطباق مدل‌ها با الزامات قانونی و اخلاقی در طول زمان	ECB (۲۰۲۴); Deloitte (۲۰۲۳); ISO/IEC ۴۲۰۰۱:۲۰۲۳; Hancke, C. H., et al. (۲۰۲۴); Gholi Pour, A., et al. (۱۳۹۶)
	افزایش اعتماد ذینفعان داخلی و خارجی به خروجی سیستم‌های مبتنی بر AI	Anderson, R., & Lee, J. (۲۰۲۴); شوندی، م.، و همکاران (۱۴۰۳); Weston, H., et al. (۲۰۱۸); Raab, W. L. (۲۰۲۲); Kaplan, R. S., & Mikes, A. (۲۰۱۶)
کارایی عملیاتی و مدیریت ریسک یکپارچه	کاهش هزینه‌های مربوط به مدیریت ریسک‌های غیرمترقبه ناشی از AI	KPMG (۲۰۰۸); Bockius, H., & Gatzert, N. (۲۰۲۴); COSO (۲۰۱۷); Arena, M., et al. (۲۰۱۰); اسماعیل زاده، پ.، و همکاران(۱۳۹۹)
	بهبود کارایی در فرآیندهای تصمیم‌گیری اصلی سازمان (Non-AI Processes) به دلیل وضوح حاکمیت	COSO (۲۰۱۷); Viscelli, T. R., et al. (۲۰۱۷); Johnson, T., & Patel, R. (۲۰۲۴); Arena, M., et al. (۲۰۱۰); Rostami, M., et al. (۱۴۰۲)
	کاهش ریسک‌های سایبری و عملیاتی به واسطه سخت‌گیری‌های حاکمیتی اعمال شده بر AI	ISO/IEC ۲۷۰۰۱:۲۰۲۲; FSB (۲۰۱۴); Power, M. (۲۰۱۶); Kim, et al. (۲۰۲۰); NIST AI RMF
	افزایش سرعت در پروژه‌های نوآورانه به دلیل وجود چارچوب‌های ریسک از پیش تعریف شده	Accenture (۲۰۲۳); Trinetix (۲۰۲۵); BCG (۲۰۲۲); Arena, M., et al. (۲۰۱۰); Schein, E. H. (۱۹۸۸)
	کاهش تعداد عدم انطباق‌های کشف شده توسط ممیزان داخلی و خارجی	Power, M. (۲۰۱۶); Sinha, V. K., & Arena, M. (۲۰۲۰); ECB (۲۰۲۴); Beasely, M. S., et al. (۲۰۲۴); COSO (۲۰۰۴)
بلوغ فرهنگ ریسک سازمانی	افزایش آگاهی کارکنان در مورد ریسک‌های خاص هوش مصنوعی (افزایش امتیاز در آموزش‌ها)	Schein, E. H. (۱۹۸۸); Anderson, R., & Lee, J. (۲۰۲۴); شوندی، م.، و همکاران Miller, P. (۲۰۲۲); Park, Y. J. (۲۰۱۹); (۱۴۰۳)
	افزایش گزارش‌دهی فعالانه (Proactive Reporting) توسط کارکنان در مورد ریسک‌های ناشناخته AI	Weston, H., et al. (۲۰۱۸); Kaplan, R. S., & Mikes, A. (۲۰۱۳); Gholi Pour, A., et al. (۱۳۹۶); Streicher, B., et al. (۲۰۱۸); Miller, P. (۲۰۲۲)
	نهادینه‌سازی نگرش "مسئولیت‌پذیری به جای سرزنش" در مواجهه با شکست‌های فناورانه	Schein, E. H. (۱۹۸۸); Raab, W. L. (۲۰۲۲); Kaplan, R. S., & Mikes, A. (۲۰۱۳); Anderson, R., & Lee, J. (۲۰۲۴); Weston, H., et al. (۲۰۱۸)
	همسویی بیشتر میان سطح تحمل ریسک اعلام شده (Appetite) و ریسک‌های عملیاتی تجربه شده (Incurred Risk)	COSO (۲۰۲۰); Bockius, H., & Gatzert, N. (۲۰۲۴); Rostami, M., et al. (۱۴۰۲); Kaplan, R. S., & Mikes, A. (۲۰۱۶); Streicher, B., et al. (۲۰۱۸)
	بهبود نتایج حاصل از فرهنگ اخلاقی در طراحی محصول (Ethical by Design)	Huang, C., et al. (۲۰۲۳); Morley, J., et al. (۲۰۲۱); FICO (۲۰۲۲); ISO/IEC ۴۲۰۰۱:۲۰۲۳; NIST AI RMF
تأثیر بر مزیت رقابتی و عملکرد کلی	توانایی سازمان در بهره‌برداری ایمن از AI برای ایجاد محصولات یا خدمات جدید (نوآوری پایدار)	Accenture (۲۰۲۳); BCG (۲۰۲۲); Trinetix (۲۰۲۵); Johnson, R. (۲۰۲۴); Mollaei, L. (۲۰۲۴)
	بهبود در بازدهی سرمایه (ROE) و کارایی کلی ناشی از اتوماسیون ریسک‌مدار	KPMG (۲۰۰۸); Hancke, C. H., et al. (۲۰۲۴); Bockius, H., & Gatzert, N. (۲۰۲۴); Johnson, R. (۲۰۲۴); Statista (۲۰۲۳)
	افزایش رتبه اعتباری و اعتبار در نزد سرمایه‌گذاران به دلیل مدیریت ریسک پیشرو	Beasley, M. S., et al. (۲۰۲۴); Power, M., et al. (۲۰۲۴); FSB (۲۰۱۴); Hancke, C. H., et al. (۲۰۲۴); Kim, et al. (۲۰۲۰)

	کاهش شکاف بین استراتژی اعلام شده و اجرای عملیاتی ریسک‌پذیری	Kaplan, R. S., & Mikes, A. (۲۰۱۶); COSO (۲۰۲۰); Viscelli, T.R., et al. (۲۰۱۷); Rostami, M., et al. (۱۴۰۲); Power, M. (۲۰۱۶)
	ارتقاء جایگاه سازمان در اکوسیستم صنعت (Industry Benchmarking) به عنوان یک بازیگر مسئول	World Economic Forum (۲۰۲۵); Accenture (۲۰۲۳); BCG (۲۰۲۲); ECB (۲۰۲۴); MIT (۲۰۲۴)

نتایج حاصل از نهادینه‌سازی فرهنگ ریسک در به‌کارگیری هوش مصنوعی، در چهار حوزه اصلی قابل مشاهده و اندازه‌گیری است: نخست، اثربخشی و قابلیت اطمینان مدل‌های **AI** که با کاهش نرخ خطا، افزایش شفافیت**(XAI)** ، کوتاه‌تر شدن زمان اصلاح مشکلات و بهبود انطباق قانونی، اعتماد ذینفعان را جلب می‌کند؛ دوم، کارایی عملیاتی و مدیریت ریسک یکپارچه که از طریق کاهش هزینه‌های ریسک‌های غیرمترقبه، ادغام ریسک‌های سیستمی، و افزایش سرعت نوآوری به واسطه چارچوب‌های از پیش تعریف‌شده ریسک، حاصل می‌شود؛ سوم، بلوغ فرهنگ ریسک سازمانی که با افزایش آگاهی، گزارش‌دهی فعالانه کارکنان در مورد شکست‌های احتمالی و نهادینه‌سازی فرهنگ "مسئولیت‌پذیری به جای سرزنش" سنجیده می‌شود؛ و در نهایت، تأثیر این اقدامات بر مزیت رقابتی و عملکرد کلی است که منجر به توانایی بهره‌برداری ایمن از نوآوری‌ها، بهبود شاخص‌های مالی **(ROE)** و ارتقاء جایگاه سازمانی در مقایسه با سایر بازیگران صنعت می‌گردد.

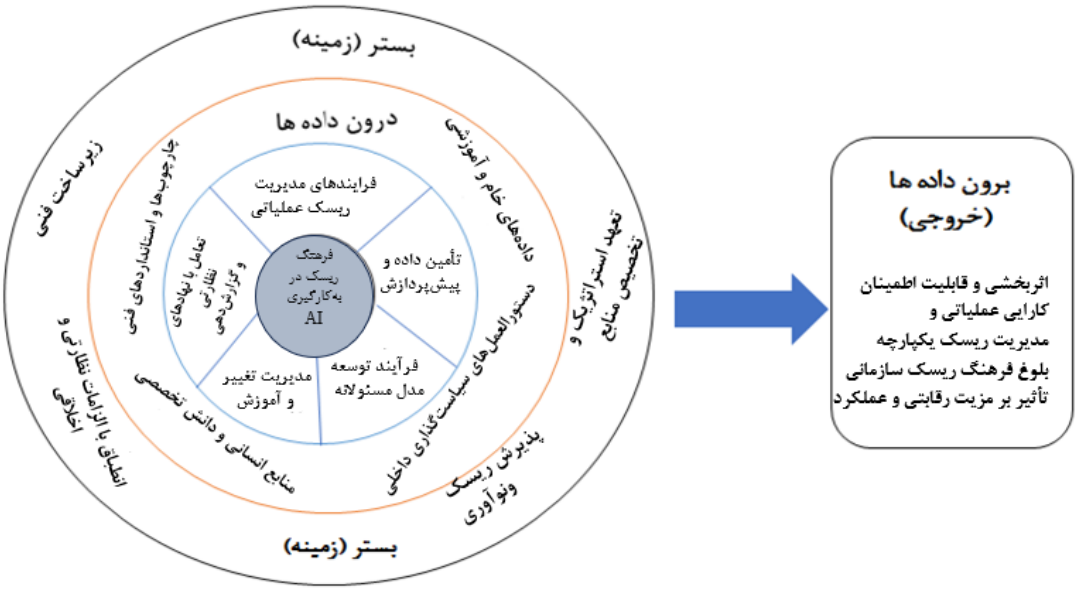
گام ششم: کنترل کیفیت

در این پژوهش به منظور بررسی روایی، از ابزار ارزیابی حیاتی گلین استفاده شده است .با کمک این ابزار، تمامی پژوهشهای منتخب به وسیله ۱۰معیار ارزشیابی و انتخاب شدند. همچنین پژوهشگر برای بررسی پایایی پژوهش، از روش توافق بین دو کدگذار بهره برده است. برای این منظور، نمونه‌ای از مقاله های برگزیدشده در اختیار خبره دیگر قرار گرفت و نتایج به دست آمده از طریق شاخص کاپا به کمک نرم افزار**SPSS** محاسبه شد که با توجه به ضریب کاپا ۰/۶۳۵ و عدد معناداری ۰/۰۰۱ پذیرش شد.

گام هفتم: ارائهٔ یافته‌ها

اساس بررسی پژوهش‌های پیشین و طبقه بندی کدهای استخراج شده در جدول‌های ۱ تا ۴، مؤلفه‌های فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری بدین صورت به دست آمد: الگوی نهایی فراترکیب شامل ۱۸ مفهوم محوری است که در چهار بخش اصلی سازماندهی شده‌اند؛ در بخش بستر و فرایندها هر کدام با ۵ مفهوم، و در بخش‌های درون‌دادها و نتایج با ۴ مفهوم سهم داشته‌اند، و این ۱۸ مفهوم در مجموع به استخراج ۸۰ کد منجر شده‌اند که توزیع این کدها در تمامی چهار بخش برابر و مساوی با ۲۰ کد برای هر بخش بوده است. با نگاهی مقایسه‌ای به ساختار مفهومی استخراج شده، درمی‌یابیم که بیشترین تمرکز کمی در این پژوهش بر روی بستر و فرایندها بوده است، زیرا هر کدام با ۵ مفهوم اصلی، بیشترین پوشش مفهومی را نسبت به دو بخش دیگر (درون‌دادها و نتایج، هر کدام ۴ مفهوم) داشته‌اند؛ با این حال، از نظر عمق کدگذاری،

توازن کامل حفظ شده است، به طوری که با وجود اختلاف در تعداد مفاهیم، هر چهار بخش اصلی مقاله (بستر، درون‌داد، فرایندها و نتایج) با استخراج دقیقاً ۲۰ کد، سهم برابری در غنای تحلیلی ۸۰ کد کلی پژوهش دارند.



شکل۳. الگوی فرهنگ ریسک به کار گیری هوش مصنوعی در صنعت بانکداری

بحث و نتیجه‌گیری

یافته‌های این پژوهش نشان داد که فرهنگ ریسک در به‌کارگیری هوش مصنوعی در صنعت بانکداری، یک سازه چندبعدی و پویا است که از تعامل چهار بُعد اصلی شامل بستر، درون‌داد، فرایند و برون‌داد شکل می‌گیرد. در این میان، بسترهای سازمانی نظیر تعهد رهبری، بلوغ حاکمیت داده و انطباق با الزامات نظارتی، نقش زیربنایی در شکل‌گیری و تقویت فرهنگ ریسک ایفا می‌کنند. این یافته با نتایج پژوهش‌های پیشین همسو است که نشان می‌دهند فرهنگ ریسک نه‌تنها یک متغیر محیطی، بلکه یک عامل تعیین‌کننده در جهت‌دهی رفتارهای سازمانی در مواجهه با ریسک‌ها است (Bockius & Gatzert, 2024; Power et al., 2024). همچنین، تأکید بر نقش تعهد مدیریت ارشد در این پژوهش با دیدگاه‌هایی که فرهنگ ریسک را به‌عنوان یک پدیده از بالا به پایین می‌دانند، همخوانی دارد (Beasley et al., 2024; Higgins, 2020). در واقع، بدون حمایت و هدایت استراتژیک مدیران ارشد، نهادینه‌سازی ارزش‌ها و هنجارهای مرتبط با مدیریت ریسک هوش مصنوعی در سطوح مختلف سازمان با چالش مواجه خواهد شد.

در بُعد درون‌دادها، یافته‌ها نشان داد که کیفیت داده‌ها، استانداردهای فنی و منابع انسانی متخصص، از جمله عوامل کلیدی در شکل‌گیری فرهنگ ریسک مؤثر هستند. این نتایج با مطالعاتی که بر اهمیت داده‌های باکیفیت و زیرساخت‌های فنی در موفقیت پروژه‌های هوش مصنوعی تأکید دارند، همراستا است (Islam et al., 2023; Mollaei, 2024). افزون بر این، نقش منابع انسانی و دانش تخصصی در این پژوهش، با دیدگاه‌هایی

که بر اهمیت سرمایه انسانی در مدیریت ریسک تأکید دارند، همخوانی دارد (Anderson & Lee, 2024; Park, 2019). به‌عبارت دیگر، حتی در حضور پیشرفته‌ترین فناوری‌ها، نبود نیروی انسانی آگاه به ریسک و فاقد درک عمیق از پیامدهای اخلاقی و عملیاتی هوش مصنوعی، می‌تواند منجر به تصمیم‌گیری‌های پرریسک و ناکارآمد شود. همچنین، توجه به منشورهای اخلاقی و سیاست‌های داخلی در این بُعد، با یافته‌های پژوهش‌هایی که بر ضرورت نهادینه‌سازی اصول اخلاقی در توسعه و استفاده از هوش مصنوعی تأکید دارند، همسو است (Huang et al., 2023; Morley et al., 2021).

در سطح فرایندی، نتایج این پژوهش بر اهمیت وجود سازوکارهای نظارتی، ارزیابی مستمر ریسک و شفافیت در تصمیمات الگوریتمی تأکید دارد. این یافته‌ها نشان می‌دهد که فرهنگ ریسک تنها در سطح نگرشی باقی نمی‌ماند، بلکه باید در قالب فرآیندهای عملیاتی و رویه‌های سازمانی تجلی یابد. این موضوع با دیدگاه‌هایی که مدیریت ریسک را یک فرآیند پویا و مستمر می‌دانند، همخوانی دارد (Fraser & Simkins, 2021; Levy et al., 2020). همچنین، تأکید بر شفافیت و قابلیت توضیح‌پذیری تصمیمات هوش مصنوعی، با پژوهش‌هایی که به چالش «جعبه سیاه» در سیستم‌های هوشمند اشاره دارند، همسو است (Chen et al., 2024; Johnson, 2024). در این راستا، ایجاد فرآیندهای MLOps و ادغام ارزیابی ریسک در چرخه عمر مدل، به‌عنوان راهکارهایی برای کاهش عدم قطعیت و افزایش کنترل‌پذیری سیستم‌های هوش مصنوعی مطرح می‌شوند. این نتایج همچنین با مطالعاتی که بر اهمیت یادگیری سازمانی و بازخورد مستمر در مدیریت ریسک تأکید دارند، همخوانی دارد (Anderson & Lee, 2024; Miller, 2022).

در بُعد برون‌دادها، یافته‌ها نشان داد که نهادینه‌سازی فرهنگ ریسک در به‌کارگیری هوش مصنوعی منجر به پیامدهایی نظیر افزایش اعتماد ذی‌نفعان، بهبود عملکرد سازمانی، کاهش هزینه‌های ناشی از ریسک و ارتقای مزیت رقابتی می‌شود. این نتایج با شواهد تجربی موجود در ادبیات مدیریت ریسک همخوانی دارد که نشان می‌دهد سازمان‌های دارای فرهنگ ریسک قوی، عملکرد بهتری در مواجهه با عدم قطعیت‌ها دارند (Beasley et al., 2024; Bockius & Gatzert, 2024). همچنین، ارتباط میان فرهنگ ریسک و اعتماد ذی‌نفعان، با یافته‌های پژوهش‌هایی که بر اهمیت شفافیت و پاسخگویی در سیستم‌های هوش مصنوعی تأکید دارند، همسو است (Curzon et al., 2021; Morley et al., 2021). از سوی دیگر، کاهش هزینه‌های ناشی از ریسک و بهبود کارایی عملیاتی، با گزارش‌های بین‌المللی که مزایای اقتصادی استفاده مسئولانه از هوش مصنوعی را نشان می‌دهند، همخوانی دارد (Statista, 2023; Trinetix, 2025).

نکته قابل‌توجه در این پژوهش، تأکید بر تعامل و هم‌افزایی میان ابعاد مختلف فرهنگ ریسک است. برخلاف برخی مطالعات که هر یک از این ابعاد را به‌صورت مجزا بررسی کرده‌اند، یافته‌های این پژوهش نشان می‌دهد که این ابعاد به‌صورت یک سیستم یکپارچه عمل می‌کنند و ضعف در هر یک از آن‌ها می‌تواند اثربخشی کل سیستم را تحت تأثیر قرار دهد. این دیدگاه با چارچوب‌های یکپارچه مدیریت ریسک سازمانی همخوانی

دارد که بر تعامل میان عوامل مختلف در مدیریت ریسک تأکید دارند (Fraser & Simkins, 2021; Sinha & Arena, 2020). همچنین، این یافته با پژوهش‌هایی که بر ماهیت سیستمی فرهنگ ریسک تأکید دارند، همسو است (Raab, 2022; Streicher, 2023). به‌عبارت دیگر، فرهنگ ریسک هوش مصنوعی را نمی‌توان صرفاً با بهبود یک مؤلفه خاص، مانند کیفیت داده یا توسعه ابزارهای فنی، تقویت کرد؛ بلکه نیازمند رویکردی جامع و چندبعدی است که تمامی جنبه‌های سازمان را دربرگیرد.

علاوه بر این، یافته‌های این پژوهش نشان می‌دهد که شکاف میان بلوغ فنی و بلوغ فرهنگی در بسیاری از سازمان‌ها، به‌ویژه در صنعت بانکداری، همچنان یک چالش اساسی است. این موضوع با گزارش‌هایی که به ناهماهنگی میان سرمایه‌گذاری‌های فناورانه و آمادگی فرهنگی سازمان‌ها اشاره دارند، همخوانی دارد (Hancke et al., 2024; Johnson & Patel, 2024). همچنین، مطالعات داخلی نیز نشان‌دهنده سطح پایین یکپارچگی فرهنگ ریسک در سازمان‌های مالی هستند که این امر می‌تواند مانعی جدی در مسیر بهره‌برداری مؤثر از هوش مصنوعی باشد (Roshani et al., 2022; Rostami Nouroozabad et al., 2023). در این زمینه، توسعه الگوهای بومی و متناسب با شرایط فرهنگی و نهادی کشورها، می‌تواند به کاهش این شکاف و ارتقای اثربخشی مدیریت ریسک کمک کند.

در مجموع، نتایج این پژوهش بر این نکته تأکید دارد که فرهنگ ریسک هوش مصنوعی در بانکداری، نه‌تنها یک ضرورت مدیریتی، بلکه یک مزیت رقابتی برای سازمان‌ها محسوب می‌شود. سازمان‌هایی که بتوانند این فرهنگ را به‌صورت مؤثر نهادینه کنند، قادر خواهند بود ضمن بهره‌گیری از فرصت‌های هوش مصنوعی، ریسک‌های مرتبط با آن را به‌صورت سیستماتیک و پایدار مدیریت نمایند. این یافته با دیدگاه‌هایی که فرهنگ ریسک را به‌عنوان یک دارایی استراتژیک در سازمان‌های دیجیتال می‌دانند، همخوانی دارد (Beasley et al., 2024; Power et al., 2024).

یکی از محدودیت‌های این پژوهش، اتکای آن به مطالعات منتشرشده به زبان‌های فارسی و انگلیسی بوده است که ممکن است منجر به نادیده گرفتن برخی منابع ارزشمند در سایر زبان‌ها شده باشد. همچنین، ماهیت کیفی روش فراترکیب و وابستگی آن به تفسیر پژوهشگر، می‌تواند بر نتایج تأثیرگذار باشد. علاوه بر این، پراکندگی مفاهیم در مطالعات اولیه و تفاوت در چارچوب‌های نظری آن‌ها، فرآیند کدگذاری و یکپارچه‌سازی یافته‌ها را با چالش مواجه کرده است. از سوی دیگر، تمرکز این پژوهش بر صنعت بانکداری، ممکن است تعمیم‌پذیری نتایج به سایر صنایع را محدود کند.

پیشنهاد می‌شود پژوهش‌های آینده با استفاده از روش‌های کمی یا ترکیبی، به اعتبارسنجی الگوی ارائه‌شده در این پژوهش بپردازند. همچنین، بررسی تطبیقی فرهنگ ریسک هوش مصنوعی در صنایع مختلف می‌تواند به درک بهتر تفاوت‌ها و شباهت‌های این مفهوم در زمینه‌های گوناگون کمک کند. استفاده از روش‌های تصمیم‌گیری چندمعیاره برای اولویت‌بندی مؤلفه‌های شناسایی‌شده نیز می‌تواند به توسعه کاربردی‌تر این الگو منجر شود. علاوه بر این، انجام مطالعات طولی برای بررسی تغییرات فرهنگ ریسک در طول زمان و در مراحل مختلف پیاده‌سازی هوش مصنوعی، می‌تواند بینش‌های ارزشمندی برای مدیران و سیاست‌گذاران فراهم آورد.

برای نهادینه‌سازی فرهنگ ریسک هوش مصنوعی در صنعت بانکداری، توصیه می‌شود سازمان‌ها اقدام به تدوین سیاست‌ها و چارچوب‌های مشخص در زمینه مدیریت ریسک این فناوری نمایند. ایجاد برنامه‌های آموزشی مستمر برای افزایش آگاهی کارکنان از ریسک‌های هوش مصنوعی، می‌تواند نقش مهمی در ارتقای فرهنگ ریسک ایفا کند. همچنین، طراحی سازوکارهای شفاف برای گزارش‌دهی و پاسخگویی در مورد تصمیمات الگوریتمی، می‌تواند به افزایش اعتماد ذی‌نفعان کمک کند. در نهایت، تقویت همکاری میان واحدهای فنی، ریسک و انطباق، و ایجاد ساختارهای حاکمیتی یکپارچه، می‌تواند زمینه‌ساز مدیریت مؤثر و پایدار ریسک‌های ناشی از هوش مصنوعی در سازمان‌ها باشد.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافعی وجود ندارد.

موازین اخلاقی

در تمامی مراحل پژوهش حاضر اصول اخلاقی مرتبط با نشر و انجام پژوهش رعایت گردیده است.

تشکر و قدردانی

از تمامی کسانی که در انجام این پژوهش ما را همراهی کردند تشکر و قدردانی به عمل می‌آید.

Extended Abstract

Introduction

The rapid expansion of artificial intelligence (AI) technologies has fundamentally transformed the operational and strategic landscape of the banking industry. Financial institutions increasingly rely on AI-driven systems for credit scoring, fraud detection, risk assessment, and customer relationship management, enabling enhanced efficiency, scalability, and decision-making precision (Samoili et al., 2020; Shavandi et al., 2025). However, alongside these benefits, AI introduces a new generation of complex, uncertain, and multidimensional risks that extend beyond traditional financial and operational domains. These risks include algorithmic bias, lack of transparency, ethical concerns, data privacy vulnerabilities, and systemic dependencies, which challenge existing risk management frameworks (Curzon et al., 2021; Islam et al., 2023). As a result, conventional risk management approaches, primarily designed for deterministic and structured systems, are often insufficient for addressing the dynamic and adaptive nature of AI technologies.

In this evolving context, the concept of risk culture has emerged as a critical determinant of organizational resilience and effectiveness in managing technological risks. Risk culture encompasses the shared values, beliefs, norms, and behaviors that shape how individuals and groups perceive, assess, and respond to risk within an organization (Bockius & Gatzert, 2024; Power et al., 2024). It represents a foundational layer that influences not only formal risk management systems but also informal decision-making processes and organizational learning mechanisms. Prior studies emphasize that many organizational failures are not due to the absence of risk management tools, but rather stem from deficiencies in risk culture, including poor communication, lack of accountability, and misalignment between strategic objectives and risk-taking behaviors (Miller, 2022; Raab, 2022). In digitally driven environments, cultivating a strong and adaptive risk culture is essential for ensuring that technological innovations are implemented responsibly and sustainably (Beasley et al., 2024; Fraser & Simkins, 2021).

With the integration of AI into banking systems, the notion of risk culture has evolved into what is increasingly referred to as AI risk culture. This concept reflects the need for organizations to internalize not only traditional risk management principles but also ethical, regulatory, and technological considerations specific to AI systems. AI risk culture requires organizations to address issues such as explainability, fairness, accountability, and continuous monitoring throughout the AI lifecycle (Chen et al., 2024; Johnson & Patel, 2024). The inherent opacity of many AI models, often described as “black box” systems, amplifies the importance of transparency and governance in maintaining stakeholder trust and regulatory compliance (Huang et al., 2023; Morley et al., 2021). Furthermore, empirical evidence suggests that organizations with more mature AI risk cultures demonstrate greater adaptability, improved risk identification capabilities, and enhanced overall performance (Anderson & Lee, 2024; Hancke et al., 2024).

Despite growing recognition of its importance, the literature on AI risk culture remains fragmented, with studies often focusing on isolated aspects such as governance, ethics, or technical risk controls. There is a lack of comprehensive frameworks that integrate the contextual, input, process, and outcome dimensions of AI risk culture in a coherent and systematic manner. Additionally, significant gaps exist between technological advancement and cultural readiness in many financial institutions, leading to increased exposure to unforeseen risks and reduced effectiveness of AI implementations (Hancke et al., 2024; Trinetix, 2025). This gap is particularly pronounced in emerging markets, where organizational risk culture maturity often lags behind technological adoption (Roshani et al., 2022; Rostami Nouroozabad et al., 2023). Therefore, there is a pressing need for research that synthesizes existing knowledge and provides a structured model to guide organizations in developing a robust AI risk culture tailored to the banking sector.

Methods and Materials

This study adopts a qualitative research design using the meta-synthesis method to systematically review and integrate findings from previous studies related to AI risk culture in the banking industry. The research population includes a broad range of academic articles, books, and dissertations published between 1988 and

2025. A comprehensive literature search was conducted across multiple databases, resulting in the identification of 145 relevant studies. Following a rigorous screening and quality assessment process using the CASP tool, 77 studies were selected for final analysis.

The analysis was conducted based on the seven-step meta-synthesis framework proposed by Sandelowski and Barroso. Data extraction involved multiple rounds of reading and coding, including open, axial, and selective coding. The extracted concepts were categorized according to a SIPOC-inspired framework consisting of four main dimensions: context (environmental and organizational conditions), inputs (resources and data), processes (operational and managerial mechanisms), and outputs (organizational outcomes). Reliability was ensured through inter-coder agreement using Cohen's kappa coefficient.

Findings

The findings reveal that AI risk culture in the banking industry is a multidimensional construct composed of four interrelated dimensions: context, inputs, processes, and outputs. Across these dimensions, a total of 18 key concepts and 80 analytical codes were identified, with each dimension contributing equally in terms of coding depth.

The contextual dimension encompasses organizational and environmental factors such as leadership commitment, data governance maturity, technological infrastructure, regulatory compliance, and organizational attitudes toward risk and innovation. These elements establish the foundational conditions necessary for embedding AI risk culture within the organization.

The input dimension includes critical resources such as high-quality training data, standardized technical frameworks, skilled human resources, and formalized internal policies. The findings indicate that the availability and quality of these inputs significantly influence the effectiveness of AI systems and the organization's ability to manage associated risks.

The process dimension highlights the operational mechanisms through which AI risk culture is enacted. These include responsible AI development practices, continuous monitoring of model performance, scenario testing, regulatory reporting, stakeholder communication, and organizational learning processes. The integration of ethical considerations and risk assessments throughout the AI lifecycle emerges as a central theme in this dimension.

The output dimension reflects the tangible and intangible outcomes of AI risk culture implementation. These outcomes include improved model accuracy and reliability, enhanced transparency and explainability, increased stakeholder trust, reduced operational risks, and strengthened competitive advantage. Additionally, organizations with mature AI risk cultures demonstrate higher levels of employee awareness, proactive risk reporting, and alignment between risk appetite and actual risk exposure.

Discussion and Conclusion

The findings of this study underscore the systemic and integrative nature of AI risk culture in the banking industry. Rather than being a static or isolated construct, AI risk culture emerges as a dynamic system shaped

by continuous interactions among contextual factors, organizational inputs, operational processes, and resulting outcomes. This holistic perspective highlights that effective risk management in AI-driven environments cannot be achieved through technical solutions alone, but requires a comprehensive alignment of organizational culture, governance structures, and operational practices.

One of the key insights of this research is the critical role of leadership and governance in establishing a strong AI risk culture. Organizational commitment at the strategic level serves as a catalyst for embedding risk-aware behaviors and ethical considerations across all levels of the organization. Without such commitment, even well-designed technical systems may fail to deliver sustainable and responsible outcomes.

Furthermore, the study emphasizes the importance of integrating ethical principles and transparency into AI processes. As AI systems increasingly influence high-stakes decisions, ensuring explainability and accountability becomes essential for maintaining trust among stakeholders. Organizations must therefore adopt proactive approaches to risk management, incorporating continuous monitoring, feedback mechanisms, and adaptive learning into their operational frameworks.

The results also highlight the significance of bridging the gap between technological advancement and cultural readiness. Many organizations invest heavily in AI technologies but neglect the cultural and organizational changes required to support their effective use. Addressing this gap requires targeted interventions, including training programs, policy development, and cross-functional collaboration.

In conclusion, this study provides a comprehensive and systematic model of AI risk culture tailored to the banking industry. By integrating multiple dimensions and synthesizing existing knowledge, the proposed framework offers valuable insights for both researchers and practitioners. It demonstrates that cultivating a robust AI risk culture is not only essential for mitigating risks but also serves as a strategic enabler for innovation, resilience, and sustainable competitive advantage in the digital era.

References

Anderson, R., & Lee, J. (2024). Organizational Learning and AI Risk Adaptation: Strategies for Dynamic Risk Management. *Journal of Business Risk Studies*, *39*(2), 145-162.

Beasley, M. S., Branson, B. C., & Hancock, B. V. (2024). Building a Risk-Intelligent Culture in Digitally-Driven Organizations. *Journal of Enterprise Risk Management*, *18*(1), 22-37.

Bockius, H., & Gatzert, N. (2024). Organizational Risk Culture: A Literature Review on Dimensions, Assessment, Value Relevance, and Improvement Levers. *European Management Journal*. https://doi.org/10.1016/j.emj.2023.12.006

Chen, L., Zhang, H., & Roberts, P. (2024). The Role of Governance in AI Risk Management: A Framework for Regulatory Compliance. *AI and Society*, *42*(1), 87-104.

Curzon, J., Kosa, T. A., Akalu, R., & El-Khatib, K. (2021). Privacy and Artificial Intelligence. *Ieee Transactions on Artificial Intelligence*, *2*(2), 96-108. https://doi.org/10.1109/TAI.2021.3088084

Esmaeilzadeh, P., Nakhjavani, S., Rad, A., & Vakili, V. (2020). *Creating and Improving Risk Culture and Risk Management Culture in Organizations with the Aim of Improving System Performance* First Scientific Research Conference on Management and Industrial Engineering,

Fraser, J., & Simkins, B. J. (2021). *Enterprise Risk Management: Today's Leading Research and Best Practices for Tomorrow's Executives*. Wiley.

Hancke, C. H., Beauperin, T., Feghali, R., & Laheurte, L. (2024). *Global Risk Manager Survey Report*.

Higgins, J. (2020). *Steps in Launching an Effective Risk Culture Program*.

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An Overview of Artificial Intelligence Ethics. *Ieee Transactions on Artificial Intelligence*, *4*(4), 799-819. https://doi.org/10.1109/TAI.2022.3194503

Islam, M., Rahman, A., & Uddin, S. (2023). AI Applications and Operational Risks in Banking: A Review. *Journal of Financial Innovation*, *7*(3), 210-227. https://doi.org/10.1007/s10203-023-00398-x

Ivanova, V. (2021). ISO 31000 - Prerequisite for Strategic Risk Management in the Activities of Organizations. *Izvestia Journal of the Union of Scientists - Varna, Economic Sciences Series*, *10*(3), 250-258. https://journals.mu-varna.bg

Johnson, R. (2024). *Risk Cultures and Banking*.

Johnson, T., & Patel, R. (2024). AI Risk Culture: A New Paradigm for Technological Governance. *Harvard business review*, *98*(4), 112-128.

Levy, C., Stegemann, U., White, O., & Helbekkmo, H. (2020). *Enterprise Risk Management and Risk Culture*. https://www.mckinsey.com/business-functions/risk/how-we-help-clients/enterprise-risk-management-and-risk-culture

Miller, P. (2022). Does Your Organization Assess Risk Culture? If Not, It Should. Here's How. *Frontiers in Research Metrics and Analytics*, *7*, 891324. https://doi.org/10.3389/frma.2022.891324

Mollaei, L. (2024). Providing a Model for Assessing the Implementation and Deployment of AI in the Banking Industry. *Frontiers in World Development and Management Journal*. https://journalfwdmj.com/index.php/fwdmj/article/download/143/140

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2021). From What to How: An Overview of AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Ai and Ethics*, *1*(1), 1-19.

Park, Y. J. (2019). *Risk Culture and Risk Management in the Australian Public Sector* https://openresearchrepository.anu.edu.au/browse?type=author&authority=1c992bca-ba2e-4b5d-b4d7-d65109e6268d

Power, M., Ashby, S., & Palermo, T. (2024). *Risk Culture in Financial Organisations*.

Raab, W. L. (2022). *The Understanding of Risk Culture Among Risk Management Practitioners* Abilene Christian University]. https://digitalcommons.acu.edu/etd/441/

Roshani, M., Rostami, N., & Badami, M. H. (2022). Developing the Dimensions of Risk Management in the Business Model. *Journal of Asset Management and Finance*, *10*(3), 67-94.

Rostami Nouroozabad, M., Badami, M. H., & Esna Ashari, M. (2023). Dimensions of Risk Management Development in the Business Model and Organizational Culture of Iran's Financial Industry. *Journal of Asset Management and Financing*, *10*(3), 67-94.

Samoili, S., Lopez Cobo, M., Gomez, E., De Prato, G., Martinez-Plumed, F., & Delipetrev, B. (2020). *AI Watch: Defining Artificial Intelligence. Towards an Operational Definition and Taxonomy of Artificial Intelligence*.

Shavandi, M., Khajeh, S., & Javidi Shahrbabak, M. J. (2025). *Trends in the Use of Artificial Intelligence in Financial Institutions and Smart Banking*.

Sinha, V. K., & Arena, M. (2020). Manifold Conceptions of the Internal Auditing of Risk Culture in the Financial Sector. *Journal of Business Ethics*, *162*(1), 81-102. https://doi.org/10.1007/s10551-018-3969-0

Statista. (2023). *Benefits of AI in Financial Services Worldwide 2023*. https://www.statista.com

Streicher, B. (2023). The Risk Culture Framework: Introducing an Integrative Framework for Holistic Risk Analysis. *Sage Open*, 1-10. https://doi.org/10.1177/21582440231191789

Trinetix. (2025). *AI in Risk Management: Building Stronger Resilience in 2025* (Trinetix Insights, Issue. https://www.trinetix.com/insights/ai-in-risk-management-building-stronger-resilience-in-2025