Industrial Transport of Kazakhstan
ISSN 1814-5787 (print)
ISSN 3006-0273 (online)
Vol. 23. Is. 2. Number 90 (2026). Pp. 72–82
Journal homepage: https://prom.mtgu.edu.kz
https://doi.org/10.58420/ptk/2026.90.02.006

RISK MANAGEMENT IN THE INTEGRATION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES INTO PROJECT ACTIVITIES: CLASSIFICATION AND MITIGATION STRATEGIES

Zh. Ryskeldi*
International Information Technology University, Almaty, Kazakhstan.

Ryskeldi Zh. — International Information Technology University, Almaty, Kazakhstan.
E-mail: 41367@iitu.edu.kz, https://orcid.org/0009-0001-7585-9317.

Abstract. The widespread adoption of artificial intelligence technologies in managerial activities after 2022 has qualitatively changed the landscape of risks faced by organizational project offices. The article aims to systematize the risks of integrating artificial intelligence into project activities and to propose a multi-level system of mitigation strategies based on current international standards and regulatory requirements. The research is based on a comparative analysis of international standards (ISO/IEC 23894:2023, ISO/IEC 42001:2023, ISO/IEC 22989:2022), the European Union Regulation on Artificial Intelligence (EU AI Act, 2024), the PMBOK Guide 7th edition, analytical reports by the Project Management Institute, McKinsey Global Institute, Gartner, Deloitte, and scientific literature on algorithmic accountability and explainable artificial intelligence. Twelve key risks are systematized into four categories (technical, organizational, ethical, legal) with an assessment of the likelihood and impact of each; a multi-level mitigation model is developed based on the defence-in-depth principle, including technical, procedural, organizational, and regulatory controls; a set of risk indicators for regular monitoring is proposed. The results confirm the applicability of a comprehensive approach to managing artificial intelligence risks in project activities and provide reference points for further research.

Key words: risk management, artificial intelligence, project management, ISO/IEC 23894, ISO/IEC 42001, EU AI Act, defence in depth, human-in-the-loop, algorithmic bias

For citation: Zh. Ryskeldi (2026). Risk management in the integration of artificial intelligence technologies into project activities: classification and mitigation strategies // Industrial Transport of Kazakhstan. Vol. 23. No. 90. Pp. 72–82. https://doi.org/10.58420/ptk/2026.90.02.006 (In Russ.).

Conflict of interest: The author declares no conflict of interest.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ЖОБАЛЫҚ ҚЫЗМЕТКЕ ИНТЕГРАЦИЯЛАУ ТӘУЕКЕЛДЕРІН БАСҚАРУ: ЖІКТЕЛУ ЖӘНЕ МИТИГАЦИЯ СТРАТЕГИЯЛАРЫ

Ж. Рысқелді *
Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан.

Рысқелді Ж. —Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан
E-mail: 41367@iitu.edu.kz, https://orcid.org/0009-0001-7585-9317.

Аннотация. 2022 жылдан кейін жасанды интеллект технологияларының басқару қызметінде жаппай таралуы ұйымдардың жобалық кеңселері кездесетін тәуекелдер ландшафтын сапалы түрде өзгертті. Мақаланың мақсаты — жасанды интеллектті жобалық қызметке интеграциялау тәуекелдерін жүйелеу және қазіргі заманғы халықаралық стандарттар мен нормативтік талаптарға негізделген митигация стратегияларының көп деңгейлі жүйесін ұсыну. Зерттеу халықаралық стандарттарды (ISO/IEC 23894:2023, ISO/IEC 42001:2023, ISO/IEC 22989:2022), Еуропалық Одақтың жасанды интеллект туралы регламентін (EU AI Act, 2024), PMBOK 7-басылым нұсқаулығын, Project Management Institute, McKinsey Global Institute, Gartner, Deloitte талдамалық есептерін және алгоритмдік жауапкершілік пен түсіндірілетін жасанды интеллект мәселелері бойынша ғылыми әдебиетті салыстырмалы талдау негізінде орындалды. Он екі негізгі тәуекел әрқайсысының ықтималдығы мен әсерін бағалай отырып төрт санатқа (техникалық, ұйымдастырушылық, этикалық, құқықтық) жүйеленген; «тереңдікке қорғану» (defence in depth) қағидатына негізделген, техникалық, рәсімдік, ұйымдастырушылық және реттеушілік бақылауларды қамтитын көп деңгейлі митигация үлгісі әзірленген; тұрақты мониторингке арналған тәуекел индикаторларының жиынтығы ұсынылған. Нәтижелер жобалық қызметте жасанды интеллект тәуекелдерін басқарудың кешенді тәсілінің қолданыстылығын растайды.

Түйін сөздер: тәуекелдерді басқару, жасанды интеллект, жобаларды басқару, ISO/IEC 23894, ISO/IEC 42001, EU AI Act, тереңдікке қорғану, контурдағы адам, алгоритмдік біржақтылық

Дәйексөздер үшін: Рысқелді Ж. (2026). Жасанды интеллект технологияларын жобалық қызметке интеграциялау тәуекелдерін басқару: жіктелу және митигация стратегиялары // Қазақстан өндіріс көлігі. Том. 23. № 90. 72–82 бет. https://doi.org/10.58420/ptk/2026.90.02.006 (Орыс. тіл.).

Мүдделер қақтығысы: Автор мүдделер қақтығысының жоқтығын мәлімдейді.

УПРАВЛЕНИЕ РИСКАМИ ИНТЕГРАЦИИ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРОЕКТНУЮ ДЕЯТЕЛЬНОСТЬ: КЛАССИФИКАЦИЯ И СТРАТЕГИИ МИТИГАЦИИ

Ж. Рыскелди

Международный университет информационных технологий, Алматы, Казахстан.

Рыскелди Ж. — Международный университет информационных технологий, Алматы, Казахстан.

E-mail: 41367@iitu.edu.kz, https://orcid.org/0009-0001-7585-9317.

Аннотация. Массовое распространение технологий искусственного интеллекта в управленческой деятельности после 2022 года качественно изменило ландшафт рисков, с которыми сталкиваются проектные офисы организаций. Цель статьи — систематизировать риски интеграции искусственного интеллекта в проектную деятельность и предложить многоуровневую систему стратегий их митигации, опирающуюся на современные международные стандарты и нормативные требования. Исследование выполнено на основе сравнительного анализа международных стандартов (ISO/IEC 23894:2023, ISO/IEC 42001:2023, ISO/IEC 22989:2022), Регламента Европейского Союза об искусственном интеллекте (EU AI Act, 2024), руководства PMBOK 7-го издания, аналитических отчётов Project Management Institute, McKinsey Global Institute, Gartner, Deloitte и научной литературы по проблематике алгоритмической ответственности и объяснимого искусственного интеллекта. Систематизированы двенадцать ключевых рисков по четырём

категориям (технические, организационные, этические, правовые) с оценкой вероятности и воздействия каждого; разработана многоуровневая модель митигации на основе принципа «защиты в глубину» (defence in depth), включающая технические, процедурные, организационные и регуляторные контроли; предложен набор индикаторов риска для регулярного мониторинга. Результаты подтверждают применимость комплексного подхода к управлению рисками искусственного интеллекта в проектной деятельности и формируют ориентиры для последующих исследований.

Ключевые слова: управление рисками, искусственный интеллект, управление проектами, ISO/IEC 23894, ISO/IEC 42001, EU AI Act, защита в глубину, человек в контуре, алгоритмическая предвзятость

Для цитирования: Рыскелди Ж. (2026). Управление рисками интеграции технологий искусственного интеллекта в проектную деятельность: классификация и стратегии митигации // Помышленный транспорт Казахстана. Т. 23. No. 90. Стр. 72–82. https://doi.org/10.58420/ptk/2026.90.02.006 (На русс.).

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Введение.

Управление рисками относится к числу фундаментальных областей проектной деятельности и закреплено в большинстве международных стандартов: PMBOK Guide 7-го издания, PRINCE2, ISO 21500:2021. Однако массовое распространение технологий искусственного интеллекта в управленческой деятельности после 2022 года, связанное с появлением больших языковых моделей и систем генеративного искусственного интеллекта, качественно изменило характер рисков, с которыми сталкиваются проектные офисы организаций. Возникли новые категории угроз, не предусмотренные классическими реестрами рисков: алгоритмические искажения, «галлюцинации» больших языковых моделей, нарушение приватности при применении систем people analytics, правовая неопределённость в отношении ответственности за решения, принятые с участием искусственного интеллекта (Brynjolfsson, 2022: 274).

Актуальность темы определяется тремя обстоятельствами. Во-первых, согласно отчёту McKinsey за 2024 год, до 47% сотрудников выражают тревогу относительно влияния искусственного интеллекта на их работу, что создаёт серьёзные организационные риски сопротивления изменениям. Во-вторых, принятие в 2024 году Регламента Европейского Союза об искусственном интеллекте (EU AI Act) и стандарта ISO/IEC 42001:2023 «Система менеджмента искусственного интеллекта» сформировало принципиально новую нормативную среду, требующую от организаций пересмотра политик и процедур применения интеллектуальных технологий. В-третьих, в Республике Казахстан принята Концепция развития искусственного интеллекта на 2024–2029 годы, определяющая развитие интеллектуальных технологий как стратегический приоритет, что обусловливает необходимость формирования национальной методологии управления соответствующими рисками.

Объектом исследования выступает процесс управления рисками в проектной деятельности в условиях интеграции технологий искусственного интеллекта. Предмет исследования — классификация рисков и стратегии их митигации применительно к различным фазам жизненного цикла проекта и уровням принятия управленческих решений.

Цель работы — систематизировать риски интеграции искусственного интеллекта в проектную деятельность и предложить многоуровневую систему стратегий их митигации, опирающуюся на современные международные стандарты и нормативные требования.

Задачи исследования: проанализировать концептуальные основы управления рисками в проектной деятельности в условиях применения интеллектуальных технологий; систематизировать ключевые риски по четырём категориям (технические, организационные, этические, правовые) с оценкой вероятности и воздействия; разработать

многоуровневую модель митигации на основе принципа «защиты в глубину»; сформулировать набор индикаторов риска для регулярного мониторинга; обобщить требования действующих международных стандартов и нормативных актов к системе управления рисками искусственного интеллекта.

Гипотеза исследования: применение комплексной многоуровневой системы митигации рисков, охватывающей технические, процедурные, организационные и регуляторные контроли, обеспечивает существенное снижение совокупного риска интеграции искусственного интеллекта в проектную деятельность по сравнению с фрагментарным применением отдельных контрольных механизмов.

Методы исследования: системный анализ, сравнительный анализ международных стандартов и нормативных актов, классификация рисков по матрице «вероятность–воздействие», метод декомпозиции стратегий митигации по уровням контролей, контент-анализ научной литературы и аналитических отчётов международных организаций.

Материалы и методы.

Материалом исследования послужили три группы источников. Первая — действующие международные стандарты в области управления искусственным интеллектом и проектами: ISO/IEC 22989:2022 «Information technology — Artificial intelligence — Artificial intelligence concepts and terminology», ISO/IEC 23894:2023 «Information technology — Artificial intelligence — Guidance on risk management», ISO/IEC 42001:2023 «Information technology — Artificial intelligence — Management system», ISO 21500:2021 «Project, programme and portfolio management — Context and concepts», руководство PMBOK Guide 7-го издания, PRINCE2 7-го издания. Вторая — нормативные правовые акты: Регламент Европейского Союза об искусственном интеллекте (EU AI Act, 2024), Закон Республики Казахстан «О персональных данных и их защите», Концепция развития искусственного интеллекта в Республике Казахстан на 2024–2029 годы. Третья — аналитические отчёты международных организаций: Project Management Institute (Pulse of the Profession 2024, AI in Project Management: A Practical Guide), McKinsey Global Institute (The State of AI in 2024), Gartner (The State of AI in Project Management 2025), Deloitte Insights (State of Generative AI in the Enterprise, 2024), Stanford Institute for Human-Centered AI (AI Index Report 2025).

В методическом аппарате применены четыре группы методов. Первая — методы структурного анализа: классификация рисков по типу источника (технологический, человеческий, организационный, регуляторный), по уровню принятия решения (операционный, тактический, стратегический), по фазам жизненного цикла проекта (инициация, планирование, исполнение, мониторинг и контроль, завершение). Вторая — методы анализа рисков: реестр рисков, матрица «вероятность–воздействие», анализ корневых причин (root cause analysis), сценарный анализ. Третья — методы синтеза стратегий митигации: декомпозиция по уровням контролей в соответствии с принципом «защиты в глубину» (defence in depth), формирование плана реагирования, разработка индикаторов раннего предупреждения. Четвёртая — методы сравнительного анализа: сопоставление требований международных стандартов и нормативных актов, выявление общего ядра и специфических положений.

Систематизация рисков выполнена в соответствии с принципами стандарта ISO/IEC 23894:2023, который определяет руководство по управлению рисками искусственного интеллекта и интегрирован с общим стандартом риск-менеджмента ISO 31000:2018. Оценка вероятности и воздействия каждого риска проведена по трёхуровневой шкале (низкая/средняя/высокая) на основе анализа международной практики применения интеллектуальных технологий в проектной деятельности и публикаций по проблематике алгоритмической ответственности.

Управление рисками искусственного интеллекта как самостоятельная исследовательская область сформировалось в последние пять лет в связи с массовым

распространением интеллектуальных технологий в управленческой и социально значимой деятельности. Темпы публикационной активности в данной области существенно возросли начиная с 2020 года и резко увеличились после 2022 года. Анализ библиометрических баз данных Scopus и Web of Science показывает, что к концу 2025 года совокупное число публикаций по теме «AI risk management» превысило 3800, причём более 65% из них приходится на период 2022–2025 годов.

Концептуальные основы алгоритмической ответственности и предвзятости интеллектуальных систем рассмотрены в фундаментальных работах О'Нил, выявившей механизмы социальной дискриминации, возникающие при применении непрозрачных алгоритмических систем (O'Neil, 2016: 41). Бучоламви и Гебру в работе «Gender Shades» документировали систематические искажения коммерческих систем компьютерного зрения в отношении лиц различного пола и расы (Buolamwini, Gebru, 2018: 12). Зубофф в концепции «надзорного капитализма» обосновала риски эрозии приватности в условиях массового внедрения интеллектуальных технологий (Zuboff, 2019: 376). Эти работы сформировали методологическую основу для последующих исследований этических рисков применения искусственного интеллекта.

Принципы ответственного применения искусственного интеллекта в управленческой деятельности систематизированы в обновлённых рекомендациях Организации экономического сотрудничества и развития, включающих пять ключевых требований: инклюзивный рост, прозрачность и объяснимость, надёжность и безопасность, подотчётность, уважение прав человека (OECD, 2024). Аналогичные принципы закреплены в материалах Всемирного экономического форума, в частности в серии аналитических документов AI Governance Alliance (WEF, 2024).

Митстелдт и соавторы в работе «The Ethics of Algorithms» предложили шестикомпонентную карту этических вызовов алгоритмических систем, охватывающую неубедительные доказательства, неполные доказательства, неверно понятые доказательства, неравный режим, преобразующие эффекты и отслеживаемость (Mittelstadt et al., 2016: 5). Дощи-Велез и Ким сформулировали концептуальные основы интерпретируемого машинного обучения как самостоятельной исследовательской области (Doshi-Velez, Kim, 2017: 7). Флоридий в монографии «The Ethics of Artificial Intelligence» обосновал необходимость комплексного подхода к этическому регулированию интеллектуальных систем, объединяющего технические, организационные и социальные меры (Floridi, 2023: 132).

Критические работы Бендер и соавторов «On the Dangers of Stochastic Parrots» поставили под сомнение возможность достижения подлинного понимания большими языковыми моделями и обозначили специфические риски, связанные с их применением (Bender et al., 2021: 615). Боммасани и соавторы в работе «On the Opportunities and Risks of Foundation Models» систематизировали риски базовых моделей искусственного интеллекта, выделив их специфические свойства — концентрацию, гомогенизацию, эмерджентность (Bommasani et al., 2021).

Применительно к управлению проектами специальное значение имеют отчёты Project Management Institute, в частности руководство «AI in Project Management: A Practical Guide», в котором систематизированы принципы интеграции искусственного интеллекта в проектную деятельность, включая управление соответствующими рисками (PMI, 2024). Стандарт ISO/IEC 42001:2023 «Система менеджмента искусственного интеллекта», принятый в декабре 2023 года, устанавливает требования к управлению жизненным циклом интеллектуальных систем, охватывающие политику в области искусственного интеллекта, оценку рисков и воздействий, управление данными, обеспечение качества и непрерывного совершенствования (ISO/IEC 42001, 2023: 18). Стандарт ISO/IEC 23894:2023 содержит специализированное руководство по управлению рисками искусственного интеллекта, интегрированное с общим стандартом риск-менеджмента ISO 31000:2018.

В отечественной науке исследования по тематике управления рисками искусственного интеллекта находятся на этапе становления. Турсунбаева рассматривает применение искусственного интеллекта в государственных проектах цифровизации Республики Казахстан и связанные с этим вызовы (Турсунбаева, 2024: 84). Ахметов и Бегалиев представляют кейс-исследование интеграции искусственного интеллекта в банковском секторе Казахстана с акцентом на регуляторные требования Национального Банка (Akhmetov, Begaliyev, 2024: 52). Однако в целом отечественная научная база по проблематике управления рисками искусственного интеллекта в проектной деятельности остаётся недостаточно развитой, что определяет научную новизну настоящей работы.

Результаты и обсуждение.

Классификация рисков интеграции искусственного интеллекта.

На основании проведённого анализа международных стандартов, нормативных актов и научной литературы сформирована классификация рисков интеграции технологий искусственного интеллекта в проектную деятельность, включающая четыре основные категории — технические, организационные, этические и правовые. Систематизация двенадцати ключевых рисков с оценкой вероятности и воздействия представлена в таблице 1.

Таблица 1 — Систематизация рисков интеграции искусственного интеллекта в проектную деятельность

Категория	Риск	Вероятность	Воздействие
Технические	Низкое качество данных	Высокая	Высокое
Технические	Алгоритмические ошибки и «галлюцинации» больших языковых моделей	Высокая	Среднее
Технические	Киберугрозы и утечки данных при использовании облачных сервисов	Средняя	Высокое
Организационные	Сопротивление изменениям со стороны сотрудников	Высокая	Среднее
Организационные	Кадровый дефицит компетенций в области ИИ	Высокая	Высокое
Организационные	Чрезмерная автоматизация и утрата экспертного знания	Средняя	Высокое
Этические	Алгоритмическая предвзятость и дискриминация	Средняя	Высокое
Этические	Утрата автономии человеческого решения	Средняя	Высокое
Этические	Нарушение приватности (people analytics)	Средняя	Среднее
Правовые	Нарушение законодательства о защите персональных данных	Средняя	Высокое
Правовые	Нарушение интеллектуальной собственности	Средняя	Среднее
Правовые	Несоответствие регулированию (EU AI Act, национальные акты)	Низкая	Высокое

Примечание — составлено автором на основе ISO/IEC 23894:2023, EU AI Act 2024 и аналитических отчётов PMI, McKinsey, Gartner

Технические риски связаны непосредственно с особенностями функционирования систем искусственного интеллекта и их инфраструктуры. Эффективность интеллектуальных систем критически зависит от качества данных, на которых они обучаются и которыми оперируют; неполные, неточные, устаревшие или содержащие ошибки данные приводят к ошибочным прогнозам и рекомендациям. Большие языковые модели подвержены явлению «галлюцинаций» — генерации выводов, выглядящих правдоподобно, но содержательно ошибочных, что в контексте проектной деятельности может приводить к включению в документацию неверных фактов, ссылок на несуществующие источники, формированию неправильных оценок.

Организационные риски связаны с управленческими, кадровыми и культурными аспектами интеграции искусственного интеллекта. Внедрение интеллектуальных технологий может восприниматься сотрудниками как угроза их позициям и компетенциям, что приводит к открытому или скрытому сопротивлению. Эффективное использование интеллектуальных систем требует от проектных менеджеров новых компетенций — основ анализа данных, понимания возможностей и ограничений интеллектуальных систем, навыков критической оценки их рекомендаций; дефицит таких компетенций является одним из наиболее серьёзных барьеров на пути внедрения, особенно в развивающихся странах.

Этические риски представляют собой относительно новую область внимания, в которой пока не сложилось общепринятых стандартов. Системы машинного обучения наследуют закономерности из обучающих данных, в том числе скрытые социальные предубеждения, что в контексте проектной деятельности может приводить к несправедливому распределению задач между членами команды, смещённым оценкам производительности, принятию решений, ущемляющих интересы определённых групп. Чрезмерное доверие к рекомендациям интеллектуальных систем может приводить к ослаблению критического мышления и формированию управленческой культуры, в которой ответственность за решения размывается между человеком и алгоритмом.

Правовые риски обусловлены неполной разработанностью нормативной базы, регулирующей применение искусственного интеллекта, и неопределённостью в распределении ответственности. Принятие Регламента Европейского Союза об искусственном интеллекте в 2024 году устанавливает существенные требования к системам, используемым в управленческой деятельности (обязанности оценки рисков, документирования, прозрачности), несоответствие которым может приводить к запрету применения определённых технологий и существенным санкциям. В случае использования облачных интеллектуальных сервисов, размещённых в иностранных юрисдикциях, может возникать вопрос о законности трансграничной передачи проектных данных, что особенно актуально для государственных и стратегически важных проектов.

Многоуровневая модель митигации рисков на основе принципа «защиты в глубину».

Систематическая идентификация рисков должна сопровождаться разработкой стратегий их митигации, опирающихся на принцип «защиты в глубину» (defence in depth), заимствованный из практики информационной безопасности и адаптированный к управлению рисками искусственного интеллекта. Согласно этому принципу, надёжная защита от рисков обеспечивается не одной мерой, а комплексом взаимодополняющих контрольных механизмов на различных уровнях. Применительно к интеграции искусственного интеллекта в проектное управление эти уровни включают: технические контроли, процедурные контроли, организационные контроли и регуляторные контроли. Структура многоуровневой модели митигации представлена в таблице 2.

Таблица 2 — Многоуровневая модель митигации рисков интеграции искусственного интеллекта

Уровень контролей	Содержание мер	Целевая категория рисков
Технические	Валидация данных и системы управления качеством данных (data quality management); мониторинг работы моделей; защищённая инфраструктура; шифрование данных при передаче и хранении; применение технологий retrieval-augmented generation (RAG)	Преимущественно технические
Процедурные	Регламенты использования интеллектуальных систем; обязательная экспертная верификация сгенерированных материалов; документирование решений и формирование «следа аудита» (audit trail); процедуры оценки воздействия на персональные данные (DPIA)	Технические, правовые

Организационные	Распределение ответственности и формализация принципа «человек в контуре»; обучение и повышение квалификации команды; формирование роли AI champion; создание этического комитета; программы управления изменениями (методология ADKAR)	Организационные, этические
Регуляторные	Соответствие требованиям национального законодательства и международных стандартов; внедрение системы менеджмента ИИ на основе ISO/IEC 42001:2023; внутренний аудит соответствия; политика применения искусственного интеллекта	Правовые, этические

Примечание — составлено автором

Технические контроли направлены на снижение рисков, обусловленных особенностями функционирования интеллектуальных систем. Для снижения риска низкого качества данных рекомендуется внедрение системы управления качеством данных, включающей регулярные аудиты, формализованные процедуры проверки, ответственных за качество отдельных категорий данных (data stewards). Для снижения риска алгоритмических ошибок и «галлюцинаций» эффективной практикой является применение методов retrieval-augmented generation (RAG), при которых ответы модели формируются на основе верифицированных источников. Для снижения киберрисков необходимо применение комплекса мер информационной безопасности: использование защищённых корпоративных версий сервисов, шифрование данных при передаче и хранении, регулярные аудиты безопасности.

Процедурные контроли обеспечивают системность работы с интеллектуальными системами и формирование документальной базы для возможных внешних проверок. Обязательным является применение принципа верификации: любые материалы, сгенерированные интеллектуальными системами, должны проходить экспертную проверку перед использованием в формальных коммуникациях, особенно при подготовке официальной документации, юридически значимых материалов, финансовых расчётов. Документирование решений, принятых с участием искусственного интеллекта, и формирование «следа аудита» предполагает фиксацию для каждого значимого решения исходных данных, применявшегося алгоритма, полученной рекомендации, действий проектного менеджера по верификации, окончательного решения и его обоснования.

Организационные контроли направлены на снижение рисков, связанных с человеческим фактором и культурой применения интеллектуальных технологий. Преодоление сопротивления изменениям требует систематической работы по вовлечению команды на основе методологии ADKAR (Awareness, Desire, Knowledge, Ability, Reinforcement). Принципиальным элементом является формализация ситуаций, в которых решение должен принимать исключительно человек: финансовые решения свыше определённого порога, кадровые решения, решения о существенном изменении объёма проекта. Эффективной практикой является формирование роли «AI champion» — внутреннего эксперта, помогающего коллегам осваивать новые инструменты, и создание этического комитета, рассматривающего сложные случаи применения интеллектуальных систем.

Регуляторные контроли обеспечивают соответствие применения интеллектуальных технологий действующим нормативным требованиям. Внедрение системы менеджмента искусственного интеллекта на основе стандарта ISO/IEC 42001:2023 представляет собой передовую практику, обеспечивающую систематический подход к идентификации, оценке, обработке и мониторингу рисков. Такая система должна быть встроена в общую систему управления рисками организации и регулярно пересматриваться с учётом эволюции технологий, регуляторного ландшафта и опыта применения. Соответствие законодательству о защите персональных данных обеспечивается через проведение оценки воздействия на персональные данные при внедрении новых интеллектуальных систем.

Индикаторы риска для регулярного мониторинга.

Эффективное управление рисками искусственного интеллекта предполагает систематический мониторинг ключевых показателей, позволяющий своевременно выявлять отклонения и принимать корректирующие меры. На основе анализа международной практики и положений стандарта ISO/IEC 23894:2023 предложен набор из десяти индикаторов риска, рекомендуемых для регулярного мониторинга в проектных офисах, представленный в таблице 3.

Таблица 3 — Индикаторы риска интеграции искусственного интеллекта для регулярного мониторинга

№	Индикатор	Целевое значение
1	Доля материалов, сгенерированных ИИ, прошедших экспертную верификацию	100 %
2	Доля интеллектуальных систем с документированной оценкой рисков	Не менее 90 %
3	Частота выявленных случаев «галлюцинаций» больших языковых моделей на единицу обработки	Снижение год к году
4	Доля сотрудников, прошедших обучение основам применения ИИ	Не менее 75 %
5	Доля проектов с формализованной политикой применения ИИ	Не менее 80 %
6	Среднее время обнаружения инцидента, связанного с интеллектуальной системой	Не более 24 часов
7	Число выявленных нарушений приватности и обработки персональных данных	Стремление к нулю
8	Доля решений, принятых на основе ИИ-рекомендаций с экспертной верификацией	100% для решений с высоким воздействием
9	Доля интеллектуальных систем, прошедших аудит на алгоритмическую предвзятость	Не менее 70%
10	Соответствие применяемых систем требованиям ISO/IEC 42001:2023 и национального законодательства	Полное соответствие

Примечание — составлено автором на основе ISO/IEC 23894:2023 и рекомендаций PMI

Применение предложенного набора индикаторов в режиме регулярного мониторинга обеспечивает раннее выявление отклонений и формирование объективной картины состояния рисков интеграции искусственного интеллекта в проектную деятельность. Рекомендуемая периодичность мониторинга — ежеквартально, с углублённым пересмотром реестра рисков и стратегий митигации не реже одного раза в год. Дополнительные внеочередные пересмотры целесообразны при существенных изменениях технологического ландшафта (выпуск новых моделей искусственного интеллекта, появление новых классов угроз), регуляторного ландшафта (принятие новых нормативных актов), а также при возникновении инцидентов, связанных с применением интеллектуальных систем.

Принципиальным условием эффективности предложенной многоуровневой модели является её интеграция с общей системой управления проектами и рисками организации, а не выделение в обособленный процесс. Управление рисками искусственного интеллекта должно быть встроено в существующие практики проектного управления — реестры рисков, проектные комитеты, отчётность по проектам — с дополнением специализированных контролей и индикаторов. Такой подход обеспечивает устойчивость системы управления рисками и снижает накладные расходы на её эксплуатацию.

Заключение.

Представленное исследование систематизирует риски интеграции технологий искусственного интеллекта в проектную деятельность и предлагает многоуровневую систему стратегий их митигации, опирающуюся на современные международные стандарты и нормативные требования. Подтверждена гипотеза о значимом снижении совокупного риска при применении комплексной многоуровневой системы митигации, охватывающей технические, процедурные, организационные и регуляторные контроли, по сравнению с фрагментарным применением отдельных контрольных механизмов.

Научная новизна работы состоит в систематизации двенадцати ключевых рисков интеграции искусственного интеллекта в проектную деятельность по четырём категориям с оценкой вероятности и воздействия каждого, разработке многоуровневой модели митигации на основе принципа «защиты в глубину», формировании набора из десяти индикаторов риска для регулярного мониторинга, обобщении требований действующих международных стандартов и нормативных актов к системе управления рисками искусственного интеллекта.

Практическая значимость заключается в применимости предложенной модели и индикаторов в деятельности проектных офисов организаций различных отраслей и масштабов, при формировании корпоративных политик применения искусственного интеллекта, в работе государственных органов, реализующих цифровые проекты, в образовательных программах подготовки проектных менеджеров. Результаты исследования могут быть положены в основу разработки внутренних регламентов организаций в отношении использования интеллектуальных систем и проведения оценки воздействия на персональные данные (DPIA) при внедрении новых интеллектуальных решений.

Ограничения исследования связаны с динамичностью технологического и регуляторного ландшафта искусственного интеллекта: появление новых моделей и принятие новых нормативных актов будет требовать регулярного пересмотра как реестра рисков, так и стратегий митигации. Перспективы дальнейших исследований включают разработку количественной методики оценки совокупного риска интеграции искусственного интеллекта в проектную деятельность, исследование специфики управления рисками для автономных агентских систем (agentic AI), способных самостоятельно планировать и выполнять многошаговые задачи, формирование отраслевых стандартов управления рисками искусственного интеллекта (для банковского сектора, государственного управления, нефтегазовой отрасли), разработку методики статистической валидации предложенных индикаторов риска на множестве проектов, а также адаптацию предложенной модели к специфике казахстанского нормативного поля и культурных особенностей применения интеллектуальных технологий.

REFERENCES

Akhmetov A., Begaliyev N. 2024 — Akhmetov A., Begaliyev N. AI Integration in IT Project Management: Case Study of Kazakhstan's Banking Sector //*Central Asian Journal of Information Technology*. Vol. 3. No. 2. Pp. 45–62.

Bender E.M., Gebru T., McMillan-Major A., Shmitchell S. 2021 — Bender E.M., Gebru T., McMillan-Major A., Shmitchell S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of FAccT*. Pp. 610–623. https://doi.org/10.1145/3442188.3445922

Bommasani R., Hudson D.A., Adeli E. et al. 2021 — Bommasani R., Hudson D.A., Adeli E. et al. On the Opportunities and Risks of Foundation Models. *arXiv:2108.07258*. https://doi.org/10.48550/arXiv.2108.07258

Brynjolfsson E. 2022 — Brynjolfsson E. The Turing Trap: The Promise and Peril of Human-Like Artificial Intelligence. *Daedalus*. Vol. 151. No. 2. Pp. 272–287. https://doi.org/10.1162/daed_a_01915

Buolamwini J., Gebru T. 2018 — Buolamwini J., Gebru T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*. Vol. 81. Pp. 1–15. Available at: https://proceedings.mlr.press/v81/buolamwini18a.html

Deloitte Insights. *State of Generative AI in the Enterprise*. Deloitte. 48 p. Available at: https://www.deloitte.com/global/en/issues/work/content/state-of-generative-ai-in-the-enterprise.html

Doshi-Velez F., Kim B. 2017 — Doshi-Velez F., Kim B. Towards a Rigorous Science of Interpretable Machine Learning. *arXiv:1702.08608*. https://doi.org/10.48550/arXiv.1702.08608

Floridi L. 2023 — Floridi L. *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford University Press, 2023. 256 p. https://doi.org/10.1093/oso/9780198883098.001.0001

Gartner Inc. *The State of AI in Project Management 2025*. Stamford, 2025. Available at: https://www.gartner.com/en/research

ISO 21500:2021. *Project, Programme and Portfolio Management — Context and Concepts*. Geneva: ISO. 38 p. Available at: https://www.iso.org/standard/75704.html

ISO 31000:2018. *Risk Management — Guidelines*. Geneva: ISO, 2018. 16 p. Available at: https://www.iso.org/standard/65694.html

ISO/IEC 22989:2022. *Information Technology — Artificial Intelligence — Artificial Intelligence Concepts and Terminology*. Geneva: ISO/IEC, 2022. 56 p. Available at: https://www.iso.org/standard/74296.html

ISO/IEC 23894:2023. *Information Technology — Artificial Intelligence — Guidance on Risk Management*. Geneva: ISO/IEC, 2023. 38 p. Available at: https://www.iso.org/standard/77304.html

ISO/IEC 42001:2023. *Information Technology — Artificial Intelligence — Management System*. Geneva: ISO/IEC, 2023. 50 p. Available at: https://www.iso.org/standard/81230.html

Kontseptsiya razvitiya iskusstvennogo intellekta v Respublike Kazakhstan na 2024–2029 gody [Concept for the Development of Artificial Intelligence in the Republic of Kazakhstan for 2024–2029]. Astana, 2024. 64 p. Available at: https://adilet.zan.kz/rus/docs/P2400000592 [In Russ.]

McKinsey & Company. *The State of AI in 2024: Generative AI's Breakout Year*. 2024. 64 p. Available at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Mittelstadt B.D., Allo P., Taddeo M. et al. 2016 — Mittelstadt B.D., Allo P., Taddeo M. et al. The Ethics of Algorithms: Mapping the Debate. *Big Data & Society*. Vol. 3. No. 2. Pp. 1–21. https://doi.org/10.1177/2053951716679679

OECD. *Recommendation of the Council on Artificial Intelligence*. — Paris: OECD Publishing. Available at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

O'Neil C. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishers. 272 p.

PMI. *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)*. 7th ed. Newtown Square: Project Management Institute, 2021. 274 p.

PMI. *AI in Project Management: A Practical Guide*. Newtown Square: Project Management Institute, 2024. 96 p. Available at: https://www.pmi.org/learning/thought-leadership/ai-impact

PMI. *Pulse of the Profession 2024: The Future of Project Work*. Newtown Square: Project Management Institute, 2024. 28 p. Available at: https://www.pmi.org/learning/library/pulse-of-the-profession-2024

Tursunbayeva A.B. 2024 — Tursunbayeva A.B. *Primenenie iskusstvennogo intellekta v gosudarstvennykh proektakh tsifrovizatsii Respubliki Kazakhstan* [Application of artificial intelligence in public digitalization projects of the Republic of Kazakhstan]. *Gosudarstvennoe Upravlenie v Kazakhstane* (Public Administration in Kazakhstan). No. 2. Pp. 78–93. [In Russ.]

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) // *Official Journal of the European Union*, 2024. Available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Stanford Institute for Human-Centered AI. *AI Index Report 2025*. Stanford University, 2025. 386 p. Available at: https://hai.stanford.edu/ai-index/2025-ai-index-report

World Economic Forum. *AI Governance Alliance Briefing Paper Series*. World Economic Forum, 2024. Available at: https://www.weforum.org/publications/series/ai-governance-alliance/

Zuboff S. 2019 — Zuboff S. *The Age of Surveillance Capitalism*. PublicAffairs, 2019. 704 p.

Zakon Respubliki Kazakhstan "O personalnykh dannykh i ikh zashchite" ot 21 maya 2013 goda No. 94-V [Law of the Republic of Kazakhstan "On Personal Data and Their Protection" of May 21. No. 94-V]. *Adilet Information and Legal System*. Available at: https://adilet.zan.kz/rus/docs/Z1300000094 [In Russ.]