

РИСКИ ВНЕДРЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В БИЗНЕС-ПРОЦЕССЫ ОРГАНИЗАЦИЙ

Г.М. Шкирятов

glebshkiryatov@yandex.ru

П.Ю. Добрыгина

dob.polia@yandex.ru

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Широкое внедрение искусственного интеллекта (ИИ) в бизнес-процессы обещает компаниям значительный рост эффективности и новые возможности. Однако использование ИИ сопряжено с рядом рисков различной природы — от технических и финансовых до юридических. В данной работе рассмотрены основные группы рисков при внедрении ИИ в деятельность организаций: технические, финансовые и правовые. Приведены примеры реальных бизнес-кейсов, иллюстрирующие указанные риски, и предложены методы их минимизации. Сделан вывод о необходимости комплексного и ответственного подхода к интеграции ИИ, включающего управление рисками для обеспечения устойчивого и безопасного развития компаний.

Ключевые слова: искусственный интеллект, авторское право, интеллектуальная собственность, нейросеть, риски, бизнес, технологии, рынок труда

Введение. Искусственный интеллект (ИИ) стремительно становится неотъемлемой частью современного бизнеса. Появление инноваций вроде ChatGPT моментально превратило ИИ, относившийся к категории тем, доступных лишь узкому кругу специалистов, в практический инструмент, получающий все более широкое распространение [1]. Компании применяют ИИ для аналитики данных, автоматизации операций, взаимодействия с клиентами и других задач, ожидая повышения производительности и конкурентных преимуществ. Однако вместе с выгодами приходят и новые угрозы. Недавние исследования и практический опыт свидетельствуют, что без должного контроля внедрение ИИ может привести к серьезным проблемам — финансовым потерям, технологическим сбоям, утечкам данных или судебным искам.

Цель данной статьи — проанализировать основные категории рисков, возникающие при интеграции ИИ в бизнес-процессы, и обсудить подходы к их снижению.

Технические риски. *Непрозрачность и сбои алгоритмов.* Сегодня работа моделей ИИ непрозрачна, т. е. это «черные ящики», которые выдают вам результат, который можно проверить только эмпирическим путем. Этот риск считается одним из наиболее опасных, поскольку эффективность решений, предлагаемых ИИ, может быть оценена

только постфактум. Его возможные ошибки в области прогнозирования спроса, стратегии продаж, оптимизации технологических процессов и т. п. могут привести к большим финансовым потерям [1]. Кроме того, сложность ИИ-систем затрудняет их отладку и может приводить к неожиданным сбоям. Показателен случай, когда автоматизированный торговый алгоритм на бирже за 45 минут ошибочной работы принес своему владельцу — компании Knight Capital — потери порядка 440 млн долл. США, едва не разорив фирму [2]. Хотя тот биржевой робот не имел элементов ИИ, пример демонстрирует, какими катастрофическими могут быть последствия неконтролируемых технических сбоев.

Утечка и компрометация данных. Функционирование ИИ-систем зависит от больших данных, включая конфиденциальную корпоративную информацию о клиентах, операциях и ноу-хау. Поэтому одной из главных технических угроз становится утечка данных при внедрении ИИ, что может стать причиной огромных штрафов и потери репутации компании. Причины утечки могут быть самыми разными — например, противоправные действия внешних лиц, сбои в работе оборудования, недобросовестность сотрудников самой компании [1]. На практике уже есть подобные инциденты. В 2023 г. инженеры Samsung Electronics случайно загрузили фрагменты секретного исходного кода в публичный чат-бот ChatGPT, что привело к разглашению служебной информации [3]. После этого компания запретила сотрудникам использовать генеративный ИИ в рабочих целях, опасаясь повторения этой ситуации. Данный случай стал показательным предупреждением для бизнеса о риске несанкционированного распространения внутренних данных через внешние ИИ-инструменты.

Алгоритмическая предвзятость. Еще один технический риск связан с тем, что ИИ-модели обучаются на исторических данных и способны унаследовать содержащиеся в них предубеждения. Алгоритмы могут начать дискриминировать определенные группы людей или искажать реальность. Известен пример Amazon, где экспериментальная система на основе ИИ для отбора персонала оказалась гендерно предвзятой — она систематически занижала рейтинг резюме женщин. Причиной стало то, что алгоритм обучился на данных о найме за предшествующие 10 лет, в течение которых в компании преобладали мужчины; модель сделала вывод, что кандидат-мужчина предпочтительнее. Несмотря на попытки откорректировать программу, гарантировать устранение скрытой дискриминации не удалось, и проект пришлось закрыть [4]. Этот кейс показывает, как неочевидные

технические особенности обучения ИИ могут привести к существенным проблемам — от упущенных ценных сотрудников до потери репутации и претензий о нарушении принципов равноправия.

Злоупотребления и кибератаки с помощью ИИ. Технологическая мощь ИИ может быть обращена и против самой организации. Злоумышленники способны использовать ИИ для автоматизации фишинговых рассылок, генерации правдоподобных ложных сообщений или даже создания подделок (deepfake). Последние представляют особую опасность: с их помощью можно имитировать голос или облик сотрудника компании для мошенничества. В Великобритании был зафиксирован случай, когда преступники с помощью нейросети создали аудиодвойник голоса генерального директора и по телефону убедили менеджера перевести им крупную сумму — около 250 тыс. долл. США [5]. Руководитель был уверен, что разговаривает со своим начальником, и не подозревал об обмане до выявления хищения. Этот инцидент демонстрирует новый тип технических рисков — угрозу целенаправленных атак на бизнес с использованием ИИ, способных обходить стандартные системы безопасности.

Наконец, к техническим рискам можно отнести *зависимость от ИИ*. Постепенно сотрудники привыкают полагаться на автоматизированные рекомендации и решения. Если внутри компании отсутствует контроль качества советов ИИ, возникает опасность утраты критического мышления и компетенций у персонала. Полная зависимость от алгоритмов без участия человека чревата тем, что в случае сбоя системы или выдачи неверного результата это сразу приведет к ошибочным действиям, а люди могут не заметить проблему вовремя. Таким образом, технические риски внедрения ИИ многогранны — от чисто инженерных проблем надежности до последствий неправильного использования технологий.

Финансовые риски. Инвестиции в решения на базе ИИ должны окупаться, однако на практике финансовые ожидания часто не оправдываются. Существует значительный *риск экономических потерь* при неудачном внедрении ИИ. Так, по оценкам RAND Corporation, более 80 % проектов в сфере ИИ оказываются неудачными, что вдвое выше среднего уровня провалов в технологических стартапах [6]. Это означает, что компании могут потратить существенные средства на разработку алгоритмов, инфраструктуру, данные и обучение персонала, но не получить обещанного эффекта. Причины разнообразны: иногда технология не достигает требуемого уровня точности, иногда организационные процессы оказываются неготовыми к изменениям, а в ряде слу-

чаев проблема заключается в завышенных ожиданиях руководства относительно возможностей ИИ [6]. Когда проект проваливается, инвестиции фактически сгорают, что напрямую сказывается на финансовых показателях компании.

Яркий пример финансового риска предоставляет кейс американской компании Zillow, работавшей на рынке недвижимости. Компания пыталась автоматизировать массовую покупку и перепродажу домов с помощью алгоритма оценки цен (сервиса Zestimate). Предполагалось, что ИИ-модель будет точно предсказывать стоимость жилья и позволять с выгодой реализовывать объекты. Однако в 2021 г. данный эксперимент обернулся крупным провалом: алгоритмы оценки недвижимости оказались ненадежными на быстро меняющемся рынке и прогнозы цен на купленные дома не оправдались. Компания была вынуждена срочно свернуть программу iBuying и распродавать приобретенные дома ниже их изначальной стоимости, фиксируя убытки. По данным СМИ, акции компании упали более чем на 30 % [7]. Этот случай стал уроком для всего сектора: даже продвинутые ИИ-алгоритмы могут ошибаться, особенно в условиях резких рыночных изменений, и такие ошибки непосредственно бьют по финансовым результатам бизнеса.

Серьезным последствием неправильно функционирующего ИИ могут быть *прямые финансовые убытки от ошибок*. Например, если алгоритм управления закупками или ценообразованием даст сбой, это приведет либо к дефициту товара, либо к перепроизводству, либо к неверным ценам. В промышленности ИИ-контроллер может вывести из строя оборудование, вызвав его простой и дорогостоящий ремонт. В сфере финансов автоматизированная торговая система, как уже упоминалось, способна в считанные минуты нанести многомиллионный ущерб компании, совершив серию неверных сделок. Таким образом, помимо затрат на разработку и внедрение у организации возникает риск неожиданных убытков от ошибочных действий ИИ.

Следует также учитывать *опосредованные финансовые риски*, связанные с социальными и регуляторными последствиями внедрения ИИ. Так, использование ИИ для оптимизации может привести к сокращению рабочих мест. Хотя снижение затрат на персонал рассматривается как экономия, массовые увольнения способны вызвать негативную реакцию общества и властей [1]. Компания рискует потерять больше денег на урегулирование конфликтов, штрафы и восстановление репутации, чем сэкономила благодаря автоматизации. Таким образом, финансовые риски охватывают не только непосредственные

убытки от сбоев, но и более широкие эффекты — от провала инвестиций в ИИ до косвенных затрат, вызванных непродуманным внедрением технологий.

Юридические риски. Правовые аспекты применения ИИ все еще формируются, и организации, внедряющие эти технологии, сталкиваются с множеством неопределенностей. *Риск нарушения законодательства и регуляторных требований* при использовании ИИ весьма высок. Одно из главных направлений — это несоблюдение норм о защите данных и приватности. ИИ-системы часто обрабатывают персональные данные клиентов, сотрудников или партнеров. Например, в Евросоюзе действует Общий регламент по защите данных (GDPR), предусматривающий штрафы до 4 % годового оборота компании за утечку персональной информации [8]. В 2023 г. ряд стран, включая Италию, временно блокировали работу ChatGPT из-за подозрений в нарушении правил конфиденциальности, потребовав от разработчиков усилить защиту данных пользователей [3]. Бизнесу необходимо учитывать подобные прецеденты: если корпоративный ИИ-сервис обрабатывает личные данные вопреки закону, компанию ждут разбирательства и санкции.

Отдельно стоит *риск дискриминации и ответственности за решения ИИ*. Если алгоритмы используют данные, содержащие предвзятость, их выводы могут нарушать антидискриминационное законодательство. В частности, ИИ, помогающий принимать решения о найме, кредитовании или страховании, может оказаться пристрастным по признаку расы, пола, возраста и т. д. В Нью-Йорке вступил в действие закон, обязывающий компании ежегодно проверять используемые рекрутинговые алгоритмы на отсутствие расовой и гендерной дискриминации. Работодатели должны публиковать отчеты о справедливости своих программ [9]. Эта норма — один из первых примеров регуляторного ответа на риски ИИ. Если компания внедрит алгоритм без аудита на беспристрастность и в результате окажется, что ИИ системно ущемляет права кандидатов или сотрудников, организацию могут ожидать коллективные иски и меры со стороны надзорных органов. Кроме того, с точки зрения действующего права ответственность за решения, принятые с участием ИИ, все равно несет компания и ее должностные лица. Это создает риск убытков от исков пострадавших, даже если непосредственной причиной ошибки был сбой нейросети.

Еще одна зона — *нарушение прав интеллектуальной собственности*. ИИ-модели обучаются на больших массивах чужих данных — текстов, изображений, аудио, зачастую защищенных авторским пра-

вом. Возникает опасность, что использование таких данных без разрешения правообладателей повлечет судебные претензии. Прецеденты уже есть: в 2024 г. восемь крупных американских газет (включая New York Daily News, Chicago Tribune и др.) подали коллективный иск против OpenAI и Microsoft, обвинив их в незаконном использовании миллионов охраняемых авторским правом статей при обучении своих ИИ-систем [10]. Издатели заявили, что технологии фактически «похитили» их контент для коммерческой выгоды, не выплатив вознаграждение журналистам. Похожий иск ранее подала The New York Times, оценивая ущерб от действий ИИ-компаний в миллиарды долларов [10]. Таким образом, внедряя ИИ, компании должны учитывать риск того, что их инструменты могли быть обучены на чьем-то контенте нелегально, либо что создаваемые ими самим ИИ-продукты нарушают чьи-то права. Помимо этого нерешенным остается вопрос, кому принадлежат права на результаты, созданные ИИ, — компании, разработчику алгоритма, или эти результаты вообще не подлежат охране как произведения, созданные не человеком. Эта правовая неопределенность также может стать основанием для споров.

Совокупность рассмотренных правовых рисков показывает, что игнорировать юридическую сторону внедрения ИИ недопустимо — штрафы, иски и запреты могут свести на нет все выгоды от инноваций. Поэтому на этапе планирования и внедрения ИИ-решений необходим юридический аудит и продуманная стратегия соответствия требованиям.

Методы минимизации рисков. Для успешной и безопасной интеграции ИИ в бизнес-процессы требуется системный подход к управлению рисками. Ниже перечислены основные методы, которые позволяют организациям минимизировать рассмотренные выше угрозы.

Тщательное планирование и пилотирование проектов ИИ. Перед широким внедрением технологии важно реалистично оценить ее возможности и ограничения. Необходимо выравнивать ожидания менеджмента с реальностью, опираясь на данные исследований и опыт других компаний. Рекомендуется начинать с небольших пилотных проектов, где можно протестировать работу ИИ в ограниченном масштабе и выявить слабые места. По результатам тестирования принимают обоснованное решение о масштабировании. Такой подход помогает избежать ситуации, когда большие средства вложены, а отдачи нет.

Обеспечение качества данных и кибербезопасности. Поскольку качество исходных данных напрямую влияет на результаты ИИ, ком-

пании должны уделять особое внимание сбору и обработке данных. Необходимо выстроить надежную систему контроля и обновлять массивы актуальной информацией. Одновременно критична защита данных: следует ограничить доступ к чувствительной информации, шифровать хранилища, отслеживать аномалии в трафике. Весь персонал, работающий с данными и ИИ, должен соблюдать строгие протоколы безопасности, чтобы снизить риск утечек. При использовании внешних облачных ИИ-сервисов нужно внимательно изучать их пользовательские соглашения и избегать передачи секретных сведений сторонним платформам. Важно помнить, что загруженные в общедоступную нейросеть данные хранятся на внешних серверах и могут стать достоянием третьих лиц, а удаление «слитой» информации затруднено [3]. По этой причине не следует использовать публичные ИИ-сервисы для работы с критически важной корпоративной информацией.

Контроль прозрачности и точности алгоритмов. Чтобы уменьшить эффект «черного ящика», важно привлекать экспертов предметной области для сверки выводов ИИ с реальностью. Например, перед применением модели в финансах ее следует проверить на исторических данных, удостоверившись в отсутствии непредвиденных убытков. В процессе эксплуатации полезно вести фиксацию решений ИИ и отслеживать ключевые метрики качества — это поможет заметить сбой в работе модели.

Обучение сотрудников и внутренние регламенты. Человеческий фактор остается центральным даже в высокотехнологичных процессах. Компаниям необходимо обучать персонал основам работы с ИИ-системами, акцентируя внимание на потенциальных рисках. Каждый сотрудник должен понимать, какие данные можно загружать в ИИ, а какие строго запрещены. Следует официально закрепить политики приемлемого использования ИИ: регламентировать, для каких задач разрешено применять нейросети, как проверять полученные от них результаты, какие меры безопасности соблюдать. Нельзя игнорировать тот факт, что сотрудники все равно захотят использовать новые инструменты — задача работодателя предвосхитить это и заранее установить границы дозволенного [11].

Юридическое сопровождение и соответствие регуляциям. Неотъемлемая часть управления рисками — работа на опережение в правовом поле. Организациям следует отслеживать изменения законодательства в области ИИ и персональных данных, консультироваться с юристами перед запуском новых ИИ-проектов. Выполнение существующих стандартов и рекомендаций способствует снижению рисков.

К примеру, — стандарт NIST-AI-600-1 от 26 июля 2024 г. «Структура управления рисками искусственного интеллекта», разработанный Национальным институтом стандартов и технологий (NIST) в соответствии с указом от 30 октября 2023 г. [1]. Придерживаясь подобных ориентиров, компания демонстрирует добросовестность и повышает готовность к будущим проверкам.

Человеческий надзор за критическими решениями. Ни при каких условиях бизнес не должен полностью отдавать принятие жизненно важных решений нейросети. Вопросы, затрагивающие здоровье людей, большие суммы денег или стратегию компании, должны решаться с участием ответственных сотрудников. Такой «гибридный» подход — связка «ИИ + человек» — позволяет воспользоваться скоростью и мощностью алгоритмов, но финальную оценку и контроль оставляет за экспертом. Это страхует от ситуаций, когда бездушный алгоритм допускает грубую ошибку, ИИ не обладает правосубъектностью и не понесет наказания за вред, причиненный его действиями, — в случае ошибок отвечать придется конкретным организациям и людям [11].

Использование надежных ИИ-платформ и аудит сторонних решений. Важно помнить, что общедоступные версии нейросетей подвержены рискам кибербезопасности, поскольку могут быть взломаны и использованы злоумышленниками для создания вредоносного кода. Бизнес-лицензии, как правило, обеспечивают регулярные обновления, современные механизмы защиты и поддержку, что снижает технологические риски [11]. Перед использованием стороннего ИИ необходимо проводить аудит: соответствует ли нейросеть требованиям компании по безопасности, где хранятся данные, как реализована защита от утечек. В договоры с вендорами ИИ имеет смысл включать положения об ответственности за инциденты (например, за компрометацию данных по вине их системы).

Реализуя перечисленные меры в комплексе, организация формирует многоуровневую систему управления рисками при работе с ИИ. Такой подход существенно повышает вероятность того, что проекты искусственного интеллекта принесут бизнесу пользу, а не проблемы.

Заключение. Внедрение искусственного интеллекта в бизнес-процессы — это шаг, открывающий перед компаниями новые горизонты эффективности и инноваций. Вместе с тем он сопряжен со значительными рисками, которые нельзя оставлять без внимания. Проведенный анализ показал, что технические риски (например, непрозрачность алгоритмов, сбои, утечки данных), финансовые риски (убытки от ошибок, некупаемые инвестиции, побочные эффекты оптимизации) и

юридические риски (несоблюдение законодательства, нарушение прав, юридическая ответственность за действия ИИ) являются реальными угрозами для организаций. Подтверждают это и рассмотренные бизнес-кейсы: от сбоя алгоритма, стоившего сотни миллионов долларов, до судебных исков СМИ к разработчикам ИИ за нарушение авторских прав.

Однако риски ИИ можно держать под контролем. Ключевой вывод состоит в том, что при ответственном и продуманном подходе компании способны значительно снизить вероятность негативных последствий. Необходимо сочетать технические меры (обеспечение качества данных, кибербезопасность, тестирование моделей) с организационными (обучение персонала, внутренние регламенты) и правовыми (соблюдение норм, диалог с регуляторами). Формируется новая дисциплина — управление рисками ИИ, и игнорировать ее в современных условиях равносильно пренебрежению техникой безопасности на производстве.

Таким образом, успешное использование ИИ в бизнесе требует баланса между смелой инновацией и осторожным риск-менеджментом. Организации, которые заранее распознают возможные угрозы и готовятся к ним, в конечном счете выиграют: они смогут извлечь из технологий ИИ максимум выгоды, минимизируя издержки и защищая себя от неприятных сюрпризов. Искусственный интеллект продолжит развиваться и внедряться во все новые отрасли, а значит, работа по обеспечению его безопасной интеграции станет неотъемлемой частью стратегии устойчивого бизнеса.

Шкирятов Глеб Михайлович — студент кафедры «Бизнес-информатика», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Добрыгина Полина Юрьевна — студентка кафедры «Бизнес-информатика», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Сусов Р.В., канд. экономических наук, доцент кафедры «Бизнес-информатика», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация. SPIN-код: 2926-4858. E-mail: susovroman@mail.ru

Ссылку на эту статью просим оформлять следующим образом:

Шкирятов Г. М., Добрыгина П. Ю. Риски внедрения искусственного интеллекта в бизнес-процессы организаций. *Политехнический молодежный журнал*, 2025, № 02 (103). URL: https://ptsj.bmstu.ru/catalog/icec/inf_tech/1102.html

THE RISKS OF INTRODUCING ARTIFICIAL INTELLIGENCE INTO THE BUSINESS PROCESSES OF ORGANIZATIONS

G.M. Shkiryatov

glebshkiryatov@yandex.ru

P.Y. Dobrygina

dob.polia@yandex.ru

Bauman Moscow State Technical University, Moscow, Russian Federation

The widespread introduction of artificial intelligence (AI) into business processes promises companies significant efficiency gains and new opportunities. However, the use of AI involves a number of risks of various nature, from technical and financial to legal. In this paper, the main groups of risks in the implementation of AI in the activities of organizations are considered: technical, financial and legal. Examples of real business cases illustrating these risks are given, and methods for minimizing them are proposed. It is concluded that there is a need for an integrated and responsible approach to AI integration, including risk management, to ensure the sustainable and safe development of companies.

Keywords: artificial intelligence, copyright, intellectual property, neural network, risks, business, technology, labor market

Received 04.07.2025

Shkiryatov G.M. — Student of Department of Business Informatics, Bauman Moscow State Technical University, Moscow, Russian Federation.

Dobrygina P.Y. — Student of Department of Business Informatics, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Susov R.V., Ph. D. (Economy), Associate Professor, Department of Business Informatics, Bauman Moscow State Technical University, Moscow, Russian Federation. SPIN-code: 2926-4858. E-mail: susovroman@mail.ru

Please cite this article in English as:

Shkiryatov G.M., Dobrygina P.Y. The risks of introducing artificial intelligence into the business processes of organizations. *Politekhnikheskiy molodezhnyy zhurnal*, 2025, no. 02 (103). URL: https://ptsj.bmstu.ru/catalog/icec/inf_tech/1102.html