

РИСКИ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ФИНАНСОВОМ ПОСРЕДНИЧЕСТВЕ И ПОДХОДЫ К ИХ ОЦЕНКЕ

Т. Н. Зверькова

Оренбургский государственный университет,
Оренбург, Россия

Предмет исследования – влияние генеративного искусственного интеллекта (GenAI) на финансовое посредничество и связанные с этим риски. **Цель исследования** – оценить текущее состояние внедрения GenAI в финансовое посредничество, определить потенциальные риски и проблемы, а также изучить последствия более широкого использования этих технологий. **Актуальность исследования** обусловлена быстрым развитием технологий искусственного интеллекта (ИИ) и их возрастающим влиянием на финансовое посредничество, что требует всестороннего анализа и понимания возможных рисков и преимуществ. **Методология исследования** основана на изучении существующих исследований по данной тематике, выявлении текущих подходов к определению рисков, эмпирическом анализе, включающем сбор и обработку перечней рисков применения GenAI и связанных с ними последствиями. Сделан **вывод**, что вариативность дефиниций в оценке рисков GenAI может создать значительные препятствия для согласованности и координации действий между различными участниками рынка. Более того, отсутствие унифицированных оценочных критериев может привести к недооценке или переоценке потенциальных угроз, что увеличивает вероятность системных сбоев и повышает риски для устойчивости системы финансового посредничества в целом.

В рамках подготовки исследования были использованы ответы, сгенерированные ChatGPT об оценке рисков, которые создает GenAI.

Ключевые слова: генеративный искусственный интеллект, GenAI, финансовое посредничество, риски, банк.

Введение

Современные достижения в области искусственного интеллекта (далее – ИИ) и, в частности генеративного (далее – GenAI), представляют собой важный этап в эволюции финансовых технологий, способных значительно автоматизировать процесс принятия решений, оптимизировать работу с большими объемами данных и улучшить взаимодействие с клиентами через создание персонализированного контента и рекомендаций. Технологические инновации в области GenAI, такие как большие языковые модели и усовершенствованные нейронные сети, позволяют обрабатывать и интерпретировать информацию с беспрецедентной скоростью и точностью. Это открывает новые возможности для анализа рынков, прогнозирования трендов и управления рисками.

Актуальность исследования определяется стремительным развитием и внедрением технологий искусственного интеллекта в сектор финансового посредничества, что приобретает особое значение в условиях растущей неопределенности и высокой вероятности возникновения непредсказуемых

последствий. Применение ИИ в финансовой сфере существенно трансформирует методы управления рисками, принятия решений и взаимодействия с клиентами, что создает как новые возможности для повышения эффективности, так и риски, связанные с технологическими сбоями, киберугрозами и этическими дилеммами. В случае с GenAI эти риски усугубляются способностью технологий генерировать контент, который может быть использован для манипуляций, мошенничества и подрыва доверия к кредитным организациям. Кроме того, актуальность данного исследования также подтверждается тем фактом, что регуляторная база и методы управления рисками часто отстают от технологических инноваций.

Целью исследования является изучение и оценка рисков, связанных с использованием GenAI в финансовом посредничестве. Исследование направлено на всестороннее понимание потенциальных рисков и неопределенностей, которые могут возникнуть при интеграции этих технологий в финансовую отрасль. Задачи исследования заключаются в определении и систематизации

специфических рисков, которые возникают при интеграции генеративного ИИ в банковский сектор.

Методология исследования включает в себя несколько этапов. Во-первых, проводится общий обзор литературы, охватывающий современные исследования и кейсы внедрения генеративного ИИ в сектор финансового посредничества. Это позволяет выявить текущие подходы к управлению рисками и определить пробелы в существующих знаниях. Вторым этапом является эмпирический анализ, включающий сбор и обработку перечней рисков применения генеративного ИИ и связанных с ними последствиями. Также проводится «экспертное интервьюирование ChatGPT», что позволяет получить дополнительную качественную оценку рисков с позиций самого генеративного ИИ и по-новому взглянуть на разработку этических норм и принципов использования ИИ в операциях банков.

В исследовании рассматривается гипотеза: «предполагаем, что использование банками GenAI повышает скорость принятия решений, но одновременно создает специфические риски, которые требуют специальных мер управления ими».

Для проверки гипотезы были проанализированы существующие исследования «и задан вопрос генеративному ИИ» (в нашем случае – модели ChatGPT), с целью выяснить, как GenAI оценивает риски, которые он может нести финансовому посредничеству?

Основная часть статьи

Обзор исследований. Анализ отечественных исследований, посвященных GenAI в финансовом посредничестве, выявляет многообразие подходов и направлений, касающихся этой технологии. Значительная часть исследований сосредоточена на изучении технических аспектов и теоретических основ ИИ. Работы таких авторов, как Г. Е. Зорин [1], С. В. Ештокин [2], представляют системный подход к определению и пониманию искусственного интеллекта, охватывая широкий спектр его применений в банковской сфере. Исследователи Е. Д. Бутенко [3] и Т. К. До [4] рассматривают ИИ как сквозную технологию, оказывающую значительное влияние на доходы цифрового рынка. Этот фактор особенно важен, поскольку ИИ становится неотъемлемой частью стратегического планирования и операционной деятельности банков. Авторы оценивают эффективность применения систем ИИ для принятия управленческих решений, акцентируя внимание на минимизации человеческого фактора и объективизации процессов принятия решений.

Вопросы применения ИИ в различных направлениях банковской деятельности находят отражение в работах, посвященных внедрению чат-ботов, дистанционного банкинга и нейронных сетей. Исследования С. В. Шекшуева и Г. В. Татьянина [5], а также А. О. Ледяйкиной [6] показывают, как ИИ способствует улучшению клиентского обслуживания и оптимизации внутренних процессов. Это подтверждают также работами Н. А. Агеева [7], Е. А. Ягуповой [8], которые отмечают, что благодаря

ИИ банки могут предоставлять более персонализированные услуги, улучшая восприятие технологичности со стороны клиентов.

Прогнозирование и моделирование финансовых показателей с использованием кластеризации на базе ИИ также является важным направлением исследований. Работы Н. И. Ломакина, М. С. Марамыгина, Е. Г. Черной [и др.] [9], а также О. В. Летова [10] изучают устойчивость банков, подчеркивая значимость ИИ в прогнозировании экономических и финансовых показателей для поддержки принятия управленческих решений.

Конкурентные преимущества и роль ИИ в цифровой трансформации банковского сектора подробно рассматриваются А. И. Боталовой [11], Г. М. Бабукиным [12], С. Н. Терещенко [13]. Они отмечают, что ИИ способствует повышению конкурентоспособности банков за счет улучшения клиентского опыта и ускорения обработки данных.

Отдельное внимание уделяется анализу рисков, связанных с внедрением ИИ. Работы А. П. Мартынова [14], О. И. Ларина [15], а также Н. М. Ломарева [16] подчеркивают актуальность вопросов информационной безопасности и противодействия мошенничеству. Авторы обращают внимание на необходимость разработки механизмов для оперативного реагирования на кибератаки и несанкционированные проникновения. М. Р. Лотош и его коллеги [17], П. М. Лукичев и О. П. Чекмарев [18], а также А. И. Болонин и А. С. Асрян [19] исследуют долгосрочные риски применения ИИ, включая хвостовые и экзистенциальные риски, которые могут ограничить потенциал вмешательства человека в процессы ИИ.

Как отмечают в своем исследовании П. М. Лукичев и О. П. Чекмарев, «прямой зависимости между увеличением использования современных технологий, таких как ИИ, и экономическим развитием, как показывает практика, пока не наблюдается. Одной из причин является нарастание рисков применения моделей GenAI для экономики и для общества в целом» [18, с. 2429].

Изученные работы показывают, что в научной среде отсутствует единое мнение в отношении рисков, присущих технологиям ИИ. Несмотря на это, большинство исследований и экспертов сходятся во мнении, что риски, сопутствующие ИИ, аналогичны рискам, связанным с другими инновациями, включая технологические, экономические, этические и риски информационной безопасности.

В отличие от рассмотренных исследований, работы Организации экономического сотрудничества и развития (сокр. ОЭСР, англ. Organisation for Economic Cooperation and Development, OECD) и Международного Валютного Фонда (МВФ) более глубоко фокусируются на выявлении и анализе рисков, связанных с внедрением GenAI в финансовом секторе, а также на разработке возможных политик и стратегий управления этими рисками.

В публикации ОЭСР (OECD) «Generative Artificial Intelligence in Finance» [20] акцент сделан на систем-

ных рисках, которые могут возникнуть вследствие широкого внедрения GenAI в финансовое посредничество. Эти риски включают в себя усугубление уже существующих, таких как рыночная нестабильность и операционные риски, а также появление новых, связанных с полной автоматизацией без участия человека. Особое внимание уделяется риску непрозрачности решений, принимаемых системами GenAI, что может затруднять их контроль и корректировку.

Соображения, высказанные специалистами МВФ Г. Шабсигом и Э.-Б. Бухеруа [21], дополняют эту картину, выделяя специфические риски, присутствующие GenAI, такие как встроенная предвзятость, проблемы конфиденциальности данных, надежность и непрозрачность алгоритмов. Отдельное внимание уделяется киберугрозам и рискам систем-

ной значимости, которые могут быть усилены при использовании GenAI, учитывая его способность создавать и передавать новые каналы рисков в финансовом секторе.

Объединение подходов российских исследований, OECD и МВФ позволяет сформировать многоуровневую структуру изучения рисков. Это включает в себя микроуровень, связанный с конкретными рисками для банков, и макроуровень, который охватывает системные риски.

Чтобы понять, какие риски необходимо более подробно исследовать, проведем небольшой сравнительный анализ и объединим в таблицу перечни рисков, которые используют в своей деятельности Банк России, АРБ, OECD [20], МВФ [21] и ChatGPT (см. таблицу 1).

Таблица 1

Перечни рисков, которые используют в своей деятельности Банк России, Ассоциация российских банков, Организация экономического сотрудничества и развития, Международный валютный фонд и ChatGPT*

Банк России	Ассоциация российских банков	ОЭСР (OECD)	МВФ	ChatGPT
Риски разработки ИИ и искажения работы модели ИИ	Ложные срабатывания ПОД/ФТ	Конфиденциальность данных	Предвзятость, дискриминация, несправедливые результаты	Точность и надежность
Риски аутсорсинга	Дезинформация клиентов	«Встроенная» предвзятость	Отсутствие объяснимости	Риск концентрации
Риски в области данных и информационной безопасности	Новые типы операционных рисков	Надежность результатов	Риски, связанные с данными	Операционные риски
Этические риски и риски нарушения прав потребителей и инвесторов	Усталость клиентов от общения с ботами	Синтетические данные	Устойчивость модели и надежность	Отсутствие прозрачности
Риски конкуренции	Ошибки при идентификации	Объяснимость результатов	Прочие риски	Этические и правовые риски
ИИ и створ	Конкуренция роботов на торгах	Кибербезопасность	Соревнование	Проблемы с адаптацией к изменениям
Злоупотребление доминирующим положением	Целенаправленное управление поведением ИИ злоумышленниками	Финансовая стабильность	Финансовая стабильность	Галлюцинации и неверные данные
Макроэкономические риски и риски финансовой стабильности	Излишняя зависимость от новых технологий		Управление	Кибербезопасность

* *Источники:* составлено автором на основании источников [20], [21], а также:

1) Применение искусственного интеллекта на финансовом рынке. Доклад для общественных консультаций / Банк России. URL: https://www.cbr.ru/Content/Document/File/156061/Consultation_Paper_03112023.pdf.

2) Гарегин Тосунян, Президент Ассоциации российских банков, на 77-й Открытой дискуссии: «Искусственный интеллект – это вовсе не панацея для банков и бизнеса» / АРБ. Открытые дискуссии Президента АРБ. URL: https://arb.ru/arb/press-on-arb/garegin_tosunyan_na_77_y_otkrytoy_diskussii_iskusstvennyy_intellekt_eto_vovse_ne-10665733/ (дата обращения: 01.08.2024).

3) Ответ от GhatGPT на вопрос «*Какие ты несешь риски для финансового посредничества?*», который был задан в ходе настоящего исследования, представлен в графе «ChatGPT».

Эмпирический анализ рисков применения генеративного ИИ и связанных с ними последствий. Рассматривая перечисленные в таблице риски, связанные с использованием GenAI в финансовом посредничестве, можно выделить несколько общих и специфических аспектов из различных источников (Банк России, АРБ, OECD, МВФ и ChatGPT). Одним из наиболее упоминаемых рисков является *предвзятость и ограниченная объяснимость*. Практически все рассмотренные нами источники признают, что генеративный ИИ может демонстрировать предвзятость из-за качества обучающих данных или архитектуры модели. Вопрос предвзятости особенно поднимается в материалах OECD.

1. Предвзятость и ограниченная объяснимость представляют значительную проблему для систем GenAI в финансовом посредничестве, которая может возникать на различных стадиях, начиная с выбора и обработки данных для обучения модели и заканчивая алгоритмическими решениями, принимаемыми в процессе работы системы.

Встроенная предвзятость определяется как несправедливое и систематическое предпочтение компьютерных систем в отношении одних лиц или групп лиц по сравнению с другими. В системе финансового посредничества это может означать, что алгоритмы, использующие машинное обучение, могут несправедливо дискриминировать клиентов на основе таких признаков, как национальность, пол, возраст или социально-экономический статус. Такая предвзятость может привести к серьезным негативным последствиям, включая несправедливое отклонение кредитных заявок, некорректную оценку рисков или предоставление неблагоприятных условий кредитования для определенных групп населения.

GenAI может быть подвержен предвзятости, создаваемой инструментами поисковой оптимизации (SEO). Эти методы включают оптимизацию контента с учетом определенных ключевых слов, метатегов и других факторов, влияющих на алгоритмы поисковых систем. В результате информация, которая более эффективно оптимизирована для SEO, может получить непропорционально высокую видимость и воздействие на пользователей. Когда GenAI обучается на таких данных, содержащих контент, оптимизированный для SEO, возникает риск, что модель будет чаще генерировать ответы, основанные на предвзятых или коммерчески мотивированных источниках. Это может исказить представление об объективной информации и повлиять на решения, принимаемые пользователями системы.

При взаимодействии с клиентами и привлечении потребителей необходимо обеспечивать прозрачность и доверие к банку и его технологиям. Банки должны ясно и понятно информировать клиентов о том, как их данные защищаются и как технологии AI/ML и GenAI обеспечивают справедливое отношение ко всем клиентам. Банки как финансовые посредники должны будут объяснять свои решения и действия, как внутренним, так и внешним заинтересованным сторонам, включая органы пруденциального надзора. *Отсут-*

ствие объяснимости может привести к недоверию клиентов, которые не могут понять, почему их заявка была отклонена или почему их транзакция была отмечена как подозрительная.

2. Конфиденциальность данных в контексте использования систем GenAI и машинного обучения представляет собой значительную проблему, когда системы AI/ML обучаются на больших объемах данных, содержащих личную и финансовую информацию, в результате которой существует риск утечки этой информации. Даже если данные были анонимизированы перед использованием, мощные алгоритмы AI/ML могут выявлять закономерности и делать выводы, которые позволяют восстановить исходные данные или идентифицировать конкретных лиц. Это создает угрозу демаскировки анонимных данных посредством логических выводов.

Системы AI/ML могут «запоминать» информацию о людях в наборе обучающих данных даже после того, как эти данные были использованы и отброшены. Это означает, что при повторном использовании или обновлении моделей может возникнуть риск непреднамеренного раскрытия или использования старых данных.

Как и в случае традиционных систем AI/ML, GenAI может стать источником утечек данных из обучающих наборов и подвергать пользователей риску демаскировки анонимных данных через сложные методы анализа. Однако уникальность GenAI заключается в его способности создавать контент, который может непреднамеренно включать конфиденциальную информацию, запомненную из обучающих данных.

Более того, модели GenAI, которые взаимодействуют с пользователями в реальном времени, могут накапливать и использовать входные данные, предоставленные пользователями, что увеличивает риск утечки конфиденциальной информации через эти взаимодействия.

Еще одной проблемой является способность GenAI анализировать и генерировать данные из различных источников, в том числе из открытых данных в Интернете. Это может привести к ситуации, когда модели непреднамеренно объединяют публичные и частные данные, создавая контент, который нарушает конфиденциальность. Например, модели могут использовать данные из социальных сетей или других онлайн-ресурсов для улучшения своих предсказаний или рекомендаций, что может привести к использованию данных без явного согласия или без ведома пользователей.

3. Точность и надежность моделей ИИ в системе финансового посредничества требует внимания к множеству факторов, особенно когда речь идет о точности результатов моделей в условиях изменяющейся среды. Прогнозирующий характер алгоритмов ИИ и машинного обучения (AI/ML) добавляет дополнительный уровень сложности. В стабильной рыночной среде, где данные относительно постоянны, модели могут генерировать надежные сигналы, которые позволяют учитывать изменения тенденций. Однако в периоды структурных сдвигов или значительных изменений в поведенческих корреляциях, эти

модели сталкиваются с проблемами. Ранее надежные сигналы могут стать ненадежными, что приводит к снижению точности прогнозирования. Так, по заявлению главы Сбербанка Г. Грефа в результате ошибок GenAI банк потерял миллиарды рублей: «Искусственный интеллект, как правило, принимает решение в больших системах. Маленькая ошибка, закравшаяся в алгоритм, может приводить к очень большим последствиям. В нашей практике мы теряли большие деньги на этом. Из-за того, что машина совершала маленькую ошибку на больших объемах, мы теряли миллиарды рублей»¹.

Также важно учитывать фактор объяснимости моделей. Использование моделей GenAI в финансовом секторе сталкивается со специфическими вызовами, связанными с надежностью и точностью производимых ими прогнозов и решений. Природа моделей GenAI такова, что они способны создавать новый контент, опираясь на данные, полученные в ходе обучения. Это создает риск генерации правдоподобных, но неверных ответов – феномен, известный как «галлюцинация». Такие галлюцинации могут быть особенно проблематичными в области разговорных систем GenAI, которые могут усиливать и уверенно защищать ложные или неточные данные.

Галлюцинации GenAI могут возникать из-за нескольких факторов, включая качество данных обучения, алгоритмические ограничения и особенности генерирования ответов на основе вероятности. Поскольку модели GenAI обучаются на огромных массивах данных, собранных из различных источников, они могут непреднамеренно интегрировать ложные или предвзятые данные, что приводит к генерации недостоверных ответов.

4. Данные и информационная безопасность.

Как отмечает Банк России, одним из перспективных направлений развития GenAI является генерация синтетических данных, то есть искусственных данных, которые не имеют отношения к конкретному клиенту или пользователю. Данные создаются специально обученной моделью ИИ, имитирующей характеристики реальной клиентской информации и могут быть использованы для обучения других моделей ИИ. Базы синтетических данных могут стать более дешевым способом повышения прогностической способности и надежности моделей, особенно в тех случаях, когда доступ к реальным данным ограничен или требует дополнительных финансовых издержек. Кроме того, обучение алгоритмов ИИ на основе синтетических данных позволит преодолеть ряд вызовов, связанных с нарушением конфиденциальности. К 2030 году синтетические данные могут полностью заменить потребность в клиентских данных для обучения моделей ИИ².

Качество синтетических данных также может быть подвержено риску из-за недостатка валидации и тести-

рования. Синтетические данные должны быть тщательно проверены на соответствие реальным данным, чтобы обеспечить их адекватность и репрезентативность. Процесс проверки включает в себя оценку статистических свойств данных, проверку на наличие скрытых корреляций и предвзятостей, а также тестирование моделей, обученных на синтетических данных, в реальных условиях. Без тщательной валидации синтетические данные могут привести к созданию моделей, которые работают хорошо на тестовых наборах, но оказываются неэффективными или даже вредными при применении в реальных условиях.

GenAI создает новые серьезные проблемы в сфере **информационной безопасности (кибербезопасности)**. GenAI может быть использован для создания более сложных фишинговых сообщений и электронных писем, которые намного труднее распознать как мошеннические. Мошенники могут использовать GenAI для имитации корпоративного стиля общения или даже личного стиля переписки сотрудников, что приводит к увеличению случаев кражи личных данных или финансовых мошенничеств.

Текущие модели GenAI все чаще подвергаются успешным атакам Jailbreak. Эти атаки основаны на разработке наборов тщательно разработанных подсказок (последовательностей слов или предложений) для обхода правил и фильтров GenAI или даже для вставки вредоносных данных или инструкций (последнее иногда называют «атакой с внедрением подсказок»). Атаки с внедрением подсказок могут привести к утечке конфиденциальной информации клиентов, нарушению операций или даже к финансовым потерям.

5. **Макроэкономические риски и риски финансовой стабильности.** Однородность в оценках рисков и принятии кредитных решений, вызванная широким использованием ИИ, может привести к схожим стратегиям управления рисками у различных финансовых посредников.

На практике, когда большое количество банков использует аналогичные модели GenAI, существует риск синхронного принятия решений, что может вызвать усиление волатильности на рынке. Например, если большинство систем GenAI, анализируя схожие данные, приходят к выводу о необходимости продажи активов в ответ на экономический стресс, это может привести к массовой распродаже и, как следствие, обрушению цен на активы. Это явление может быть особенно опасным в периоды рыночной неопределенности, когда панические настроения могут легко распространиться и усугубить кризис. В июле 2024 г. индексы Nasdaq 100 упали более чем на 3 %, рынок обвалился на 1 трлн долларов, когда у инвесторов возникли вопросы о том, сколько времени потребуется, чтобы окупилась инвестиция в технологии GenAI. Роботы стали производить

¹ Греф рассказал о миллиардных потерях Сбербанка из-за ошибок искусственного интеллекта. Публикация от 26.02.2019 г. URL: <https://www.kommersant.ru/doc/3895309> (дата обращения: 01.08.2024).

² Применение искусственного интеллекта на финансовом рынке / Банк России. URL: https://www.cbr.ru/Content/Document/File/156061/Consultation_Paper_03112023.pdf.

распродажи пакетов акций. В лидерах обвала оказались компании из сферы ИИ, список возглавили полупроводниковые компании, такие как NVIDIA Corp., Broadcom Inc. и Arm Holdings plc³.

В условиях рыночной эйфории GenAI, анализируя настроения и поведение участников рынка, может прийти к завышенной оценке активов, что приведет к «пузырям» на финансовых рынках. Поскольку модели GenAI обучаются на исторических данных, они могут не всегда учитывать внезапные изменения рыночных условий или появление новых факторов риска, что приводит к неправильному ценообразованию активов. Это усиливает проциклический характер финансовых рынков, где положительные и отрицательные колебания усиливаются, создавая потенциальные системные риски. Так, в начале 2023 г. чат-бот от Google неправильно ответил на вопрос и обрушил акции компании. Акции Alphabet, материнской компании Google, рухнули почти на 9 % во время презентации в Париже чат-бота с искусственным интеллектом Bard⁴.

Концентрация риска на нескольких поставщиках технологий GenAI приводит к проблеме, когда банки становятся частично заложниками своих технологических партнеров, что может вынудить регулятора вмешиваться для предотвращения краха, если такой поставщик столкнется с серьезными проблемами. Это создает парадоксальную ситуацию, при которой технологические компании, хотя и не являются финансовыми посредниками, могут стать системно значимыми для финансового рынка. В случае значительного сбоя или проблем в одном из таких технологических гигантов финансовая стабильность целого сектора может оказаться под угрозой, вызывая необходимость в государственном вмешательстве для стабилизации ситуации.

Выводы

Проведенное исследование подтверждает гипотезу, что применение GenAI в финансовом посредничестве направлено на повышение эффективности процессов и улучшение принятия решений. Однако, наряду с преимуществами, возникает ряд специфических рисков, которые регулятор и исследовательские организации в основном оценивают схожим образом. Тем не менее, существенная проблема заключается в отсутствии четких и стандартизированных количественных и качественных показателей для оценки этих рисков. Отсутствие единообразия в трактовках и измерениях рисков приводит к тому, что банки могут по-разному оценивать угрозы, исходящие от применения ИИ. Такая вариативность в оценке рисков может создать значительные препятствия для согласованности

и координации действий между различными участниками рынка. В результате могут возникнуть неоднородные подходы к управлению рисками, что затруднит разработку эффективных стратегий для их минимизации. Более того, отсутствие унифицированных оценочных критериев может привести к недооценке или переоценке потенциальных угроз, что увеличивает вероятность системных сбоев и повышает риски для устойчивости финансовой системы в целом. На данную проблему обращает внимание и Банк России, который в своем докладе, обращаясь к участникам рынка, просит разрабатывать и предлагать качественные и количественные показатели для оценки значимости рисков использования ИИ на финансовом рынке⁵.

Учитывая вышеизложенное, следует подчеркнуть, что оценка значимости рисков, связанных с использованием GenAI, требует применения как качественных, так и количественных показателей для обеспечения всестороннего анализа. Качественные показатели включают в себя анализ потенциальных угроз, связанных с технологиями ИИ, таких как риски безопасности и конфиденциальности данных, а также проблемы с интерпретируемостью и объяснимостью решений, принимаемых ИИ. Количественные показатели включают в себя различные параметры, которые могут измерять степень воздействия потенциальных рисков на финансовые результаты и операционные процессы. Одним из таких показателей является частота инцидентов, связанных с нарушением безопасности или утечкой данных, когда ИИ-системы оказываются уязвимыми.

В дополнение к этим параметрам также можно использовать показатели, отражающие эффективность и точность ИИ-систем. Это включает в себя оценку точности прогнозов и решений, производимых ИИ, а также уровень автоматизации процессов и влияние на операционные затраты. Такие показатели, как средняя ошибка прогнозов или процент ошибок в принятии решений, могут дать представление о том, насколько надежны и точны используемые ИИ-системы.

Также важным показателем является время отклика или задержки, вызванные ИИ-системами, которое может повлиять на эффективность работы и качество обслуживания клиентов. Время, необходимое для обработки запросов и выполнения задач, может быть критическим для банковских операций и служить индикатором потенциальных рисков.

Наконец, для комплексной оценки рисков важно учитывать аспекты соблюдения нормативных требований и стандартов, а также способность ИИ-систем адаптироваться к изменениям в законодательстве и регуляторных нормах. Это может быть измерено

³ Индекс Nasdaq 100 обвалился на \$1 трлн в худший день с 2022 года (публикация от 25.07.24) / ФИНАМ : сайт. URL: <https://www.finam.ru/publications/item/indeks-nasdaq-100-obvalilsya-na-1-trln-v-khudshiy-den-s-2022-goda-20240725-1104/> (дата обращения: 01.08.2024).

⁴ Чат-бот от Google неправильно ответил на вопрос и обрушил акции компании (публикация от 09.02.2023) / Деловой календарь. URL: <https://officelife.media/news/40560-chat-bot-ot-google-nepravilno-otvetil-na-vopros-i-obrushil-aktsii-kompanii/> (дата обращения: 01.08.2024).

⁵ Применение искусственного интеллекта на финансовом рынке / Банк России: офиц. сайт. URL: https://www.cbr.ru/Content/Document/File/156061/Consultation_Paper_03112023.pdf.

через частоту проверок на соответствие, а также через наличие и качество документации по соблюдению нормативных требований.

По итогам проведенного исследования считаем возможным предложить для обсуждения несколько показателей и формул для оценки следующих рисков, связанных с использованием GenAI в финансовом посредничестве:

1. Предвзятость и ограниченная объяснимость.

Индекс дисперсии предвзятости (Bias Dispersion Index, или BDI) – оценивает вариацию предвзятости в различных ситуациях, помогая выявить, где предвзятость наиболее выражена.

$$BDI = \frac{1}{n} \sum_{i=1}^n (B_i - \bar{B})^2, \quad (1)$$

где B_i – уровень предвзятости для отдельного случая;
 \bar{B} – средний уровень предвзятости, n – общее число случаев.

2. Конфиденциальность данных.

Среднее время обнаружения утечки данных (Average Data Breach Detection Time, или ADT) – измеряет среднее время, которое проходит от момента утечки данных до ее обнаружения, и помогает оценить эффективность системы мониторинга безопасности:

$$ADT = \frac{\sum_{i=1}^n T_i}{n}, \quad (2)$$

где T_i – время, необходимое для обнаружения утечки данных,
 n – количество инцидентов.

3. Точность и надежность моделей ИИ:

– *средняя ошибка прогнозирования (Mean Absolute Error, или MAE)* – измеряет среднюю абсолютную разницу между фактическими и прогнозируемыми значениями, что отражает точность модели:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (3)$$

где y_i – фактическое значение,
 \hat{y}_i – прогнозируемое значение,
 n – количество наблюдений;

– *среднеквадратическая ошибка (Root Mean Squared Error, или RMSE)* – оценивает среднеквадратическое отклонение прогнозируемых значений от фактических и используется для оценки отклонений при моделировании данных:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}; \quad (4)$$

– *средняя ошибка предсказаний (MAPE)* – средний процент ошибок в предсказаниях модели по отношению к реальным значениям, измеряет среднюю абсолютную разницу между фактическими и прогнозируемыми значениями, что отражает точность модели:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \times 100. \quad (5)$$

где A_i – фактическое значение,
 F_i – прогнозируемое значение,
 n – количество наблюдений.

4. Данные и информационная безопасность.

Ущерб от утечек данных (Data Breach Impact, или DBI) – оценивает среднюю стоимость ущерба от утечек данных, включая прямые финансовые потери, штрафы, компенсации и косвенные расходы, связанные с утратой репутации и необходимостью улучшения мер безопасности:

$$DBI = \frac{\text{Общая стоимость ущерба от утечек данных}}{\text{Количество утечек данных}}. \quad (6)$$

5. Макроэкономические риски и риски финансовой стабильности:

– *индекс интерпретируемости решений (Decision Interpretability Index)* – оценка по шкале от 1 до 10, где 1 – полное отсутствие интерпретируемости, 10 – полная прозрачность решений. Оценка может быть получена путем экспертного анализа;

– *индекс системного риска (SRM)* – вероятность того, что сбои в работе ИИ-систем приведут к значительным финансовым потерям в масштабе системы:

$$SRM = \frac{\sum_{i=1}^N L_i \times P_i}{N}, \quad (7)$$

где L_i – возможные потери,
 P_i – вероятность их наступления,
 N – количество сценариев;

– *индекс мультиканальной согласованности (IMC)* – оценка влияния SEO-предвзятости и качества модели с использованием мультиканальных подходов:

$$IMC = \frac{1}{n} \sum_{i=1}^n \left| \frac{R_i - A_i}{A_i} \right| \times 100, \quad (8)$$

где n – количество каналов (например, органический поиск, платный поиск, социальные сети и т.д.),
 R_i – прогнозируемое значение (например, количество кликов, конверсий, CTR) для канала i ,
 A_i – фактическое значение для канала i .

Предложенные нами показатели и формулы могут помочь количественно и качественно оценить риски, связанные с применением ИИ в финансовом посредничестве. Использование этих показателей в комбинации позволяет создать более полное представление о значимости рисков, связанных с применением ИИ, и разработать соответствующие стратегии для их управления и минимизации.

Литература

1. Зорин Г. Е. Искусственный интеллект и его применение в банковской сфере // Вестник Российского университета кооперации. 2020. № 1 (39). С. 31–36.
2. Ештокин С. В. Использование искусственного интеллекта для развития цифровой системы современных моделей взаимодействия коммерческих банков и их клиентов // Экономика и социум: современные модели развития. 2020. Т. 10, № 4. С. 381–390. DOI: 10.18334/ecsoc.10.4.111409
3. Бутенко Е. Д. Искусственный интеллект в банках сегодня: опыт и перспективы // Дайджест-финансы. 2020. Т. 25, № 2 (254). С. 230–242. DOI: 10.24891/df.25.2.230
4. До Т. К. Применение искусственного интеллекта в управлении лояльностью клиентов в банках // Информация и инновации. 2023. Т. 18, № 2. С. 56–64. DOI: 10.31432/1994-2443-2023-18-2-56-64
5. Шекшуева С. В., Татьяна Г. В. Внедрение чат-ботов с искусственным интеллектом в дистанционное банковское обслуживание как способ повышения конкурентоспособности коммерческого банка // Современные наукоемкие технологии. Региональное приложение. 2023. № 3 (75). С. 47–51. DOI: 10.6060/snt.20237503.0006
6. Леядякина А. О., Чугунов В. И. Развитие дистанционного обслуживания в коммерческом банке на основе технологий искусственного интеллекта // Управленческий учет. 2021. № 11-2. С. 426–432. DOI: 10.25806/uu11-22021426-432
7. Агеев Н. А. Искусственный интеллект в банковской сфере: цифровизация российских банков / В сборнике: Цифровое пространство: экономика, управление, социум // Сборник научных статей V Юбилейной Всероссийской научной конференции (Смоленск, 17 мая 2023 года). – Курск: ЗАО «Университетская книга», 2023. С. 5–13.
8. Ягулова Е. А., Черникова Л. Ф., Новикова Т. Ю. Современное состояние и перспективы использования искусственного интеллекта в банках // Экономика и предпринимательство. 2023. № 11 (160). С. 1057–1060. DOI: 10.34925/EIP.2023.160.11.200
9. Кластеризация в исследовании закономерностей развития и устойчивости российских коммерческих банков системами искусственного интеллекта / Н. И. Ломакин, М. С. Марамыгин,

- Е. Г. Черная [и др.] // Фундаментальные исследования. 2024. № 2. С. 17–25. DOI: 10.17513/fr.43565
10. Летов О. В. Этические аспекты в области разработки искусственного интеллекта (Обзор) // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 3: Философия. 2024. № 1. С. 34–44. DOI: 10.31249/rphil/2024.01.03
11. Боталова А. И. Применение искусственного интеллекта в разных подразделениях коммерческого банка // Вектор экономики. 2020. № 4 (46). URL: https://elibrary.ru/download/elibrary_42807747_91626717.pdf.
12. Бабукин Г. М. Цифровизация и искусственный интеллект в банках: шаг в будущее // Chronos: экономические науки. 2021. Т. 6, № 1 (29). С. 6–9.
13. Терещенко С. Н. Цифровая трансформация бизнес-процессов коммерческих банков под влиянием технологий искусственного интеллекта // Наука Красноярья. 2020. Т. 9, № 4. С. 415–427. DOI: 10.12731/2070-7568-2020-4-415-427
14. Мартынова А. П. Применение искусственного интеллекта в банках // Общество и цивилизация. 2024. Т. 6, № 2. С. 51–54.
15. Ларина О. И., Шкробтанова С. С. Искусственный интеллект в банках: первые результаты, препятствия и риски // Региональная экономика. Юг России. 2023. Т. 11, № 4. С. 88–98. DOI: 10.15688/region.2023.4.8
16. Ломарев Н. М. Использование искусственного интеллекта в риск-менеджменте банка // Современная математика и концепции инновационного математического образования. 2021. Т. 8, № 1. С. 447–456.
17. Лотош М. Р., Платонов В. В., Ткалич П. П. Барьеры на пути внедрения искусственного интеллекта в российских банках: размеры, причины, сроки и пути преодоления // Вопросы инновационной экономики. 2021. Т. 11, № 1. С. 315–332. DOI: 10.18334/vines.11.1.111529
18. Лукичев П. М., Чекмарев О. П. Долгосрочные риски применения искусственного интеллекта в экономике // Вопросы инновационной экономики. 2023. Т. 13, № 4. С. 2427–2442. DOI: 10.18334/vines.13.4.119948
19. Болонин А. И., Асрян А. С. Риски внедрения искусственного интеллекта в банковский сектор // Инновации и инвестиции. 2024. № 2. С. 288–291.
20. Generative Artificial Intelligence in Finance – OECD: Artificial Intelligence Papers, 2023, No. 9. – 35 p. URL: <https://www.oecd-ilibrary.org/docserver/ac7149cc-en.pdf?expires=1725101672&id=id&accname=guest&checksum=0AC80487C25EFE8FB16F5410604028D3> (дата обращения: 01.08.2024).
21. Shabsigh G., Boukherouaa E.-B. Generative Artificial Intelligence in Finance: Risk Considerations. – International Monetary Fund, 2023. – 24 p. URL: <https://www.imf.org/en/Publications/fintech-notes/Issues/2023/08/18/Generative-Artificial-Intelligence-in-Finance-Risk-Considerations-537570> (дата обращения: 01.08.2024).

Сведения об авторе

Зверькова Татьяна Николаевна – кандидат экономических наук, доцент кафедры банковского дела и страхования, Оренбургский государственный университет, Оренбург, Россия.
E-mail: tnzverkova@mail.ru

RISKS OF GENERATIVE ARTIFICIAL INTELLIGENCE IN FINANCIAL INTERMEDIATION AND APPROACHES TO THEIR ASSESSMENT

T. Zver'kova
Orenburg State University,
Orenburg, Russia

The subject of the study is the impact of generative artificial intelligence (GenAI) on financial intermediation and associated risks. The purpose of the study is to assess the current state of GenAI adoption in financial intermediation, identify potential risks and challenges, and explore the implications of increased use of these technologies. The relevance of the study is due to the rapid development of artificial intelligence (AI) technologies and their increasing impact on financial intermediation, which requires a comprehensive analysis and understanding of possible risks and benefits. The research methodology is based on the study of existing research on this topic, identification of current approaches to identifying risks, empirical analysis, including collection and processing of lists of risks of using GenAI and associated consequences. It is concluded that the variability of definitions in risk assessment GenAI can create significant obstacles to consistency and coordination of actions between various market participants. Moreover, the lack of unified assessment criteria can lead to underestimation or overestimation of potential threats, which increases the likelihood of system failures and increases risks to the sustainability of the financial intermediation system as a whole.

As part of the preparation of the study, responses generated by ChatGPT about the assessment of the risks posed by GenAI were used.

Keywords: generative artificial intelligence, GenAI, financial intermediation, risks, bank.

References

1. Zorin G. E. Artificial intelligence and its application in the banking sector, *Vestnik Rossiiskogo universiteta kooperatsii*, 2020, No. 1 (39), pp. 31-36.

2. Eshtokin S. V. Using artificial intelligence for the development of a digital system of modern models of interaction between commercial banks and their clients, *Ekonomika i sotsium: sovremennye modeli razvitiya*, 2020, Vol. 10, No. 4, pp. 381–390. DOI: 10.18334/ecsoc.10.4.111409

3. Butenko E. D. Artificial intelligence in banks today: experience and prospects, *Daidzhest-finansy*, 2020, Vol. 25, No. 2 (254), pp. 230-242. DOI: 10.24891/df.25.2.230

4. Do T. K. Application of artificial intelligence in managing customer loyalty in banks, *Informatsiya i innovatsii*, 2023, Vol. 18, No. 2, pp. 56-64. DOI: 10.31432/1994-2443-2023-18-2-56-64.

5. Shekshueva S. V., Tat'yanin G. V. Introduction of chatbots with artificial intelligence into remote banking services as a way to increase the competitiveness of a commercial bank, *Sovremennye naukoemkie tekhnologii. Regional'noe prilozhenie*, 2023, No. 3 (75), pp. 47–51. DOI: 10.6060/snt.20237503.0006.

6. Ledyaikina A. O., Chugunov V. I. Development of remote servicing in a commercial bank based on

artificial intelligence technologies, *Upravlencheskii uchet*, 2021, No. 11-2, pp. 426–432. DOI: 10.25806/uu11-22021426-432.

7. Ageev N. A. Artificial intelligence in the banking sector: digitalization of Russian banks, *Tsifrovoye prostranstvo: ekonomika, upravlenie, sotsium* [Digital space: economics, management, society], Collection of scientific articles of the V Anniversary All-Russian Scientific Conference (Smolensk, May 17, 2023). Kursk: ZAO "Universitetskaya kniga", 2023, pp. 5–13.

8. Yagupova E. A., Chernikova L. F., Novikova T. Yu. Current state and prospects for the use of artificial intelligence in banks, *Ekonomika i predprinimatel'stvo*, 2023, No. 11 (160), pp. 1057–1060. DOI: 10.34925/EIP.2023.160.11.200

9. Lomakin N. I., Maramygin M. S., Chernaya E. G. et al. Clustering in the study of patterns of development and stability of Russian commercial banks with artificial intelligence systems, *Fundamental'nye issledovaniya*, 2024, No. 2, pp. 17–25. DOI: 10.17513/fr.43565

10. Letov O. V. Ethical aspects in the field of artificial intelligence development (Review), *Sotsial'nye i gumanitarnye nauki. Otechestvennaya i zarubezhnaya literatura. Ser. 3: Filosofiya*, 2024, No. 1, pp. 34–44. DOI: 10.31249/rphil/2024.01.03

11. Botalova A. I. Application of artificial intelligence in different divisions of a commercial bank, *Vektor ekonomiki*, 2020, No. 4 (46), pp. 62.

12. Babukin G. M. Digitalization and artificial intelligence in banks: a step into the future, *Chronos: ekonomicheskie nauki*, 2021, Vol. 6, No. 1 (29), pp. 6-9.

13. Tereshchenko S. N. Digital transformation of business processes of commercial banks under the influence of artificial intelligence technologies, *Nauka Krasnoyar'ya*, 2020, Vol. 9, No. 4, pp. 415-427. DOI: 10.12731/2070-7568-2020-4-415-427.

14. Martynova A. P. Application of artificial intelligence in banks, *Obshchestvo i tsivilizatsiya*, 2024, Vol. 6, No. 2, pp. 51-54.

15. Larina O. I., Shkrebtanova S. S. Artificial intelligence in banks: first results, obstacles and risks, *Regional'naya ekonomika. Yug Rossii*, 2023, Vol. 11, No. 4, pp. 88-98. DOI: 10.15688/re.volsu.2023.4.8

16. Lomarev N. M. Use of artificial intelligence in bank risk management, *Sovremennaya matematika i kontseptsii innovatsionnogo matematicheskogo obrazovaniya*, 2021, Vol. 8, No. 1, pp. 447-456.

17. Lotosh M. R., Platonov V. V., Tklich P. P. Barriers to the implementation of artificial intelligence

in Russian banks: dimensions, reasons, timing and ways to overcome, *Voprosy innovatsionnoi ekonomiki*, 2021, Vol. 11, No. 1, pp. 315-332. DOI: 10.18334/vinec.11.1.111529

18. Lukichev P. M., Chekmarev O. P. Long-term risks of using artificial intelligence in economics, *Voprosy innovatsionnoi ekonomiki*, 2023, Vol. 13, No. 4, pp. 2427-2442. DOI: 10.18334/vinec.13.4.119948.

19. Bolonin A. I., Asryan A. S. Risks of introducing artificial intelligence into the banking sector, *Innovatsii i investitsii*, 2024, No. 2, pp. 288-291.

20. Generative Artificial Intelligence in Finance, OECD Artificial Intelligence Papers, December 2023, No. 9, 35 p. Available at: <https://www.oecd-ilibrary.org/docserver/ac7149cc-en.pdf?expires=1725101672&id=id&accname=gu est&checksum=0AC80487C25EFE8FB16F5410604028D3> (Date of application: 08/01/2024).

21. Shabsigh G., Boukheroua E.-B. Generative Artificial Intelligence in Finance: Risk Considerations // *Fintech Notes*, Vol. 2023, Is. 006. Retrieved Aug 2, 2024. DOI: <https://doi.org/10.5089/9798400251092.063>

About the author

Tat'yana N. Zver'kova – PhD in Economics, Associate Professor of the Department of Banking and Insurance, Orenburg State University, Orenburg, Russia.
E-mail: tnzverkova@mail.ru

Вниманию специалистов!

Брюханова Н. В.

Теоретические и практические аспекты использования бюджетирования в финансовом планировании инвестиционной деятельности организаций / Н. В. Брюханова, Н. В. Фадейкина, М. М. Фридрих: учеб. пособие. – 2-е изд., перераб. и доп.; Новосиб. гос. ун-т экономики и управления. – Новосибирск: НГУЭУ, 2023. – 164 с.

Рассматриваются вопросы использования организациями метода бюджетирования в финансовом планировании инвестиционной деятельности с целью формирования многофакторной модели результатов, повышения качества управления инвестиционными процессами, снижения уровня неопределенности и риска. На теоретическом уровне обоснована целесообразность применения бюджетирования как одного из основных методов финансового планирования инвестиционной деятельности организаций.

Пособие адресовано студентам (магистрантам), аспирантам, научно-педагогическим работникам вузов, а также специалистам, занимающимся вопросами финансового управления и планирования, инвестиционной деятельности и инвестиционного проектирования.

