



Банк России

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА ФИНАНСОВОМ РЫНКЕ

Доклад для общественных консультаций

Москва
2023

ОГЛАВЛЕНИЕ

| | |
|---|----|
| Введение..... | 2 |
| Ключевые тезисы..... | 3 |
| 1. Определение и основные технологии искусственного интеллекта | 5 |
| 2. Глобальные тренды развития искусственного интеллекта | 11 |
| 3. Применение технологий искусственного интеллекта на финансовом рынке | 14 |
| 4. Условия развития искусственного интеллекта на российском финансовом рынке | 21 |
| 5. Риски искусственного интеллекта на финансовом рынке..... | 27 |
| 6. Подходы к регулированию искусственного интеллекта | 33 |
| 6.1. Зарубежный опыт регулирования искусственного интеллекта..... | 33 |
| 6.2. Подходы к регулированию искусственного интеллекта на российском финансовом рынке..... | 42 |
| Вопросы для общественного обсуждения | 47 |
| Глоссарий | 49 |
| Список литературы | 51 |

Доклад подготовлен совместно Департаментом финансовых технологий и Департаментом стратегического развития финансового рынка.

Ответы на вопросы, поставленные в докладе, а также замечания и предложения к нему просим направлять до 15.12.2023 включительно на электронный адрес ai_doklad@cbr.ru.

Фото на обложке: Shutterstock/FOTODOM

107016, Москва, ул. Неглинная, 12, к. В

Официальный сайт Банка России: www.cbr.ru

© Центральный банк Российской Федерации, 2023

ВВЕДЕНИЕ

Банк России продолжает серию публикаций по вопросам инноваций на финансовом рынке¹. Данный доклад посвящен искусственному интеллекту (ИИ).

Для целей доклада под ИИ понимается комплекс технологических решений, позволяющий имитировать когнитивные функции человека и при выполнении конкретных задач получать результаты, сопоставимые с результатами его интеллектуальной деятельности. Комплекс технологических решений включает в себя программное обеспечение, информационно-коммуникационную инфраструктуру, процессы и сервисы по обработке данных и поиску решений².

На сегодняшний день ИИ уже широко используется участниками российского финансового рынка в различных бизнес-процессах: для взаимодействия с клиентами, управления рисками, аналитики, мониторинга и совершения операций и так далее.

При этом потенциал дальнейшего расширения использования ИИ финансовыми организациями представляется значительным. ИИ может повышать эффективность финансовых организаций и качество оказываемых ими услуг.

Тем не менее интенсивное развитие технологий ИИ, особенно в части генеративного ИИ³, создает не только новые возможности, но и несет определенные вызовы и риски, в связи с чем, с одной стороны, необходимо создать условия для развития технологии ИИ, а с другой – определить направления развития регулирования.

Доклад посвящен вопросам развития и применения ИИ на финансовом рынке⁴, включая определение ИИ, его типы, основные технологии ИИ и тенденции их развития. Также рассматриваются вопросы применения ИИ в отдельных секторах финансового рынка, условия развития технологий ИИ и возможные риски внедрения ИИ финансовыми организациями. Отдельное внимание уделено обзору мировых тенденций регулирования ИИ. Также сформулирован подход к регулированию ИИ на российском финансовом рынке.

Ключевые вопросы для обсуждения приведены в конце доклада. Ответы на вопросы, замечания и предложения к докладу просим направлять до 15 декабря 2023 года включительно на электронный адрес ai_doklad@cbr.ru.

¹ Ранее Банк России опубликовал серию докладов по вопросам развития современных цифровых инноваций в финансовой сфере – в частности, доклады для общественных консультаций [«Цифровой рубль»](#), [«Криптовалюты: тренды, риски, меры»](#), [«Развитие рынка цифровых активов в Российской Федерации»](#), аналитические доклады [«Децентрализованные финансы»](#), [«Токенизированные безналичные деньги на счетах в банках»](#).

² Определение ИИ из Указа Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации».

³ Тип ИИ, который может создавать новый контент (текст, изображение, аудио и видео) в ответ на запросы пользователя, что является ключевым отличием от традиционного ИИ.

⁴ Доклад не затрагивает вопросы использования ИИ Банком России в своей деятельности.

КЛЮЧЕВЫЕ ТЕЗИСЫ

1. Искусственный интеллект представляет собой особым образом настроенные технологические решения, которые позволяют имитировать когнитивные функции человека и получать результаты, сопоставимые с реализацией его интеллектуальной деятельности.

Ключевое отличие систем ИИ от обычных программных средств заключается в способности ИИ к обучению. Системы на основе ИИ обучаются, извлекая закономерности из данных и настраивая внутренние скрытые параметры, необходимые для получения решения.

2. Основным трендом развития ИИ на ближайшие годы становится широкое распространение генеративного ИИ, в том числе больших языковых моделей, в то время как ранее внимание уделялось в основном традиционному ИИ.
3. Финансовый сектор вошел в число сфер экономики, в которых технологии ИИ внедряются наиболее активно⁵.

ИИ постепенно становится технологией общего назначения и способен трансформировать бизнес-процессы и рыночные практики финансовых организаций. Внедрение ИИ может способствовать росту эффективности бизнеса участников рынка, в том числе за счет снижения издержек, ускорения процессов, ресурсной оптимизации, обработки больших массивов данных, а также повышения лояльности клиентов, благодаря более качественным и таргетированным предложениям продуктов и сервисов.

Среди ключевых направлений применения ИИ на финансовом рынке можно выделить скоринг, андеррайтинг, торговлю, инвестиционное консультирование, управление рисками, маркетинг, взаимодействие с клиентами, противодействие мошенничеству и прочее. Причем перечень сфер применения ИИ постоянно расширяется. С учетом экспансии ИИ, роста глубины интеграции данной технологии в бизнес-процессы финансовых организаций он может оказать существенное влияние на развитие российского финансового рынка.

4. Дальнейшая реализация потенциала технологии ИИ на российском финансовом рынке зависит от повышения доступности и качества данных, развития и повышения доступности программного обеспечения и вычислительной инфраструктуры, развития научной базы в области ИИ, развития профессиональных компетенций в области ИИ, обеспечения доверия граждан к технологиям ИИ, развития конкуренции, а также от создания регуляторных условий, способствующих развитию технологии ИИ с учетом возникающих рисков.
5. В мире отсутствует единое мнение в отношении рисков, присущих технологии ИИ. Вместе с тем большинство экспертов склоняется к тому, что риски, характерные для ИИ, во многом схожи с рисками для других технологий. В частности, выделяются риски технологического, экономического, этического характера, а также риски в сфере информационной безопасности. При этом эксперты отмечают наличие специфических рисков, связанных с применением технологий ИИ, особенно генеративного ИИ.
6. На данный момент в мировой практике регулирования ИИ условно можно выделить три основных подхода: ограничительный, гибридный и стимулирующий.
7. В отношении регулирования технологии ИИ в России стоит отметить следующее.
 - Банк России считает целесообразным поддерживать создание условий, направленных на стимулирование развития ИИ на финансовом рынке с учетом риск-ориентированного принципа регулирования ИИ.
 - Широкое распространение использования генеративного ИИ и расширение сфер использования технологии ИИ в целом на финансовом рынке может быть сопряжено с появлением

⁵ [Artificial Intelligence Index Report](#). Stanford University (2023).

нием новых, специфических для данной технологии рисков, требующих дополнительного анализа и оценки. Выявление факторов, значимых для определения необходимости регуляторного вмешательства, будет осуществляться в рамках систематического мониторинга Банка России.

- Банк России в целом придерживается технологически нейтрального подхода к регулированию, позволяющему финансовым организациям внедрять новые технологии и развивать ИИ, и на данный момент не видит необходимости в оперативной разработке отдельного регулирования использования технологии ИИ финансовыми организациями.
- Внедрение специальных требований, ориентированных на ИИ, может быть рассмотрено в отдельных случаях после проведения дополнительных консультаций с участниками рынка и с их возможной проработкой в рамках механизма пилотирования цифровых инноваций (в том числе экспериментальных правовых режимов) с учетом риск-факторов, указанных в докладе.
- При этом Банк России отмечает важность самостоятельного управления финансовыми организациями рисками, связанными с применением ИИ, и выстраивания в финансовых организациях, применяющих ИИ, системы управления такими рисками, учитывающей характер и масштаб таких рисков, а также особенности применяемых программ на основе ИИ. Кроме того, приветствуется обмен лучшими практиками в сфере ИИ между участниками рынка.
- Дополнительно по инициативе финансовых организаций возможно привлечение Банка России к участию в наблюдении за проектами по внедрению ИИ и анализе связанных с ними рисков.

1. ОПРЕДЕЛЕНИЕ И ОСНОВНЫЕ ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Определение искусственного интеллекта

История ИИ началась в 1950-х годах с появлением первых электронных вычислительных машин и развитием экспертных систем⁶. На протяжении многих лет ИИ оставался предметом научного интереса, а системы ИИ создавались прежде всего исследовательскими институтами. Перспективы массового коммерческого применения большинства технологий ИИ были ограничены в силу недостаточного объема вычислительных мощностей, а также отсутствия требуемого объема данных. В 2000-х и 2010-х годах увеличение скорости и памяти компьютеров, а также увеличение объемов данных вследствие развития и роста проникновения Интернета позволили преодолеть прежние барьеры для развития ИИ⁷.

В настоящий момент существует большое количество определений ИИ, поскольку активное внедрение продуктов и сервисов на его основе в самых разных отраслях меняет смысловое наполнение термина, что затрудняет создание исчерпывающего и достаточно гибкого определения, которое охватывает все подходы к реализации ИИ. Кроме того, ИИ – это широкая область знаний, которая опирается на исследования разных наук (компьютерных наук, статистики, экономики, нейробиологии, лингвистики, психологии и философии), каждая из которых имеет свой понятийный аппарат, предметную область исследований и методологическую базу.

Для целей данного доклада предлагается следующее определение: **ИИ – это комплекс технологических решений, позволяющий имитировать когнитивные функции человека и при выполнении конкретных задач получать результаты, сопоставимые с результатами его интеллектуальной деятельности.** Комплекс технологических решений включает в себя программное обеспечение, информационно-коммуникационную инфраструктуру, процессы и сервисы по обработке данных и поиску решений⁸.

Категории искусственного интеллекта

Ряд экспертов выделяют три категории ИИ на основе возможностей технологии: **слабый ИИ, сильный ИИ и суперсильный ИИ**⁹.

Слабый ИИ (Artificial Narrow Intelligence, ANI – ИИ узкого назначения) определяется как набор алгоритмов, ориентированный на решение узкоспециализированных, прикладных задач.

В рамках слабого ИИ условно можно выделить три типа:

- **Традиционный ИИ** – тип ИИ, имитирующий умственную деятельность и способный делать прогнозы и выводы на основе анализа. В финансовой отрасли используется для кредитного скоринга, борьбы с мошенничеством, управления рисками и других задач.
- **Адаптивный ИИ**¹⁰ – тип ИИ, в рамках которого модель может обучаться и динамически адаптироваться к изменениям в окружающей среде за счет сбора обратной связи в режиме реального времени¹¹.

⁶ Система, имитирующая экспертные решения на основе заранее составленной для нее базы знаний в соответствующей области.

⁷ [History of Artificial Intelligence](#). Harvard (2017).

⁸ Определение ИИ из Указа Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации».

⁹ [What is strong AI? IBM](#); [Artificial Intelligence – automotive's new value-creating engine](#). McKinsey (January 2018); [NASA Framework for the Ethical Use of Artificial Intelligence \(AI\)](#). NASA (April 2021).

¹⁰ [Gartner – Top Strategic Technology Trends 2023](#). Gartner (October 2022).

¹¹ [Искусственный интеллект – основа для создания финансовых услуг нового поколения](#). Ассоциация ФинТех (Август 2023).

- **Генеративный ИИ** – тип ИИ, который может создавать новый контент (текст, изображение, аудио и видео) в ответ на запросы пользователя, что является ключевым отличием от традиционного ИИ.

При этом растущие возможности генеративного ИИ, а также его способность выполнять более широкий по сравнению с традиционными моделями ИИ спектр задач рассматриваются некоторыми исследователями как шаг к эволюции традиционных алгоритмов ИИ в сторону так называемого «сильного ИИ»¹².

Сильный ИИ (Artificial General Intelligence, AGI – ИИ общего назначения) – гипотетическая концепция системы ИИ, способной к решению любого класса задач, которые могут быть выполнены человеком¹³. Среди возможных свойств сильного ИИ выделяют способность к **автономному принятию решений и самостоятельной постановке задач**, построению логических умозаключений, стратегическому планированию и действиям в условиях значительной неопределенности, а также возможность самообучения¹⁴.

Национальная стратегия развития искусственного интеллекта на период до 2030 года¹⁵ указывает на важность фундаментальных научных исследований в области сильного ИИ и подчеркивает возможность возникновения как позитивных, так и негативных последствий его появления. Появление моделей сильного ИИ прогнозируется рядом экспертов не ранее 2030–2040-х годов. При этом некоторые ученые считают, что создание сильного ИИ невозможно, поскольку мышление человека неалгоритмично и часто имеет невербализованный характер, который не может воспринять программа¹⁶.

Суперсильный ИИ (Artificial Superintelligence, ASI – супер-ИИ) – концепция, выделяемая некоторыми исследователями и описывающая возможность создания в отдаленной перспективе систем ИИ, которые будут **существенно превосходить разум человека**. Перспектива появления суперсильного ИИ является неопределенной, а контуры и смысловое содержание этой концепции на сегодня не сформированы. Суперсильный ИИ в большей степени остается предметом футуристических рассуждений.

Технологии искусственного интеллекта: краткий обзор

Ключевое отличие технологий ИИ от обычных программных средств заключается в способности ИИ к обучению. Обычные информационные системы предполагают заложенный программистом и алгоритмом способ достижения результата, а **системы ИИ обучаются** на основе данных и заранее описанных правил обучения, извлекая закономерности из данных и настраивая внутренние скрытые параметры, необходимые для получения решения¹⁷.

1. **Машинное обучение (Machine Learning, ML) – технология ИИ**, подразумевающая использование математических моделей, которые на основе данных помогают компьютеру обучаться без непосредственных инструкций со стороны человека. **Машинное обучение используется как в традиционном, так и в генеративном искусственном интеллекте**. Применение инструментов машинного обучения позволяет выявлять скрытые, неочевидные для человека взаимосвязи в выборке данных и формировать на их основе прогнозы и рекомендации.

¹² [Initial policy considerations for generative artificial intelligence](#). OECD (September 2023); [Generative Artificial Intelligence: toward a new civilization?](#) Arthur D. Little (October 2023).

¹³ Райков А.Н. Слабый vs Сильный искусственный интеллект // Информатизация и связь. 2020. № 1.

¹⁴ [Каким бывает искусственный интеллект?](#) ПАО Сбербанк (Ноябрь 2022).

¹⁵ Утверждена Указом Президента Российской Федерации от 10.10.2019 № 490.

¹⁶ [Каким бывает искусственный интеллект?](#) ПАО Сбербанк (Ноябрь 2022).

¹⁷ Emilia Bonaccorsi di Patti, Filippo Calabresi, Biagio De Varti, Fabrizio Federico, Massimiliano Affinito, Marco Antolini, Francesco Lorizzo, Sabina Marchetti, Ilaria Masiani, Mirko Moscatelli, Francesco Privitera and Giovanni Rinna. [No. 721 – Artificial intelligence in credit scoring: an analysis of some experiences in the Italian financial system](#). Bank of Italy (October 2022).

Машинное обучение является самой востребованной технологией ИИ. Согласно исследованию, проведенному Ассоциацией ФинТех, около 90% методов и инструментов ИИ, которые используют компании-респонденты, основаны на методах машинного обучения¹⁸.

Основные способы машинного обучения

- **Обучение с учителем (Supervised Learning)** – вид машинного обучения, при котором обучение модели происходит на входных данных, предварительно размеченных для решения конкретной задачи. Разметка данных – процесс добавления меток в сырые данные, нацеленный на то, чтобы показать модели целевые ответы, которые она должна предсказывать. Данный вид обучения в том числе используется для *кредитного скоринга*.
- **Обучение без учителя (Unsupervised Learning)** – вид машинного обучения, при котором модель не получает заранее размеченных данных, а выявляет закономерности между ними самостоятельно. Типичный пример обучения без учителя – обнаружение аномалий, например, мониторинг транзакций для ПОД/ФТ.
- **Обучение с подкреплением (Reinforcement Learning)** – модель обучается во взаимодействии со средой, оптимизируя свое поведение на основе получаемого за различные действия вознаграждения. Данный вид машинного обучения позволяет адаптироваться к изменяющейся среде, а также моделировать большое количество различных сценариев работы системы, учитывая поведение других участников рынка.

В числе задач, для решения которых используются алгоритмы машинного обучения, можно выделить следующие:

- **Задача регрессии** – прогнозирование на основе данных, обладающих различными признаками. Классический пример решения этой задачи в финансовой сфере – *кредитный скоринг клиента* на основе различных типов финансовых (история банковских транзакций, кредитная история и налоговые декларации) и нефинансовых (возраст клиента, его образование, место работы) данных.
- **Задача классификации** – определение класса объекта на основе признаков. Результат обучения алгоритма – классификатор, который прогнозирует класс для новых экземпляров данных. Пример – *классификация ценных бумаг по риску и предполагаемой доходности с целью оптимизации портфеля*.
- **Задача кластеризации** – распределение по группам отдельных экземпляров данных на основании наличия у них схожих признаков. Кластеризацию можно также использовать для определения в наборе данных связей, которые сложно логически отследить просмотром или наблюдением сведений. Финансовые компании решают данную задачу при помощи алгоритмов ИИ – например, при *сегментировании клиентов*.
- **Задача выявления аномалий** – выявление аномальных значений в выборке стандартных данных. На финансовом рынке наиболее распространенным примером решения данной задачи при помощи алгоритмов ИИ является *отслеживание аномалий в транзакциях пользователей и борьба с мошенничеством (антифрод)*.

К числу преимуществ использования алгоритмов классического машинного обучения эксперты относят их более **высокую объяснимость и меньшие требования к времени обучения**. Простота и относительно низкая стоимость данного вида обучения позволяет применять его для тестирования гипотез. Вместе с тем специфика математических моделей, лежащих в основе систем классического машинного обучения, позволяет использовать его для решения ограниченного круга задач.

¹⁸ [Искусственный интеллект – основа для создания финансовых услуг нового поколения](#). Ассоциация ФинТех (Август 2023).

2. **Глубокое обучение (Deep Learning)** – вид машинного обучения на основе **искусственных нейронных сетей (Artificial Neural Networks)** – моделей ИИ, которые представляют собой многослойную систему соединенных и взаимодействующих логических вычислительных единиц (искусственных нейронов). Целью создания искусственных нейронных сетей была попытка имитировать процесс человеческого мышления. Модель строится по тем же принципам, что и нервная система человека. Так, генеративная модель ChatGPT компании OpenAI и YandexGPT компании «Яндекс», формирующие текстовые ответы на запросы пользователей, основаны на модели глубокого обучения Generative Pre-trained Transformer (GPT).

Искусственная нейронная сеть содержит слои взаимосвязанных искусственных нейронов: входной слой, скрытые слои и выходной слой. Информация поступает из входного слоя, узлы в котором обрабатывают данные и передают на скрытые слои. Каждый скрытый слой анализирует выходные данные предыдущего слоя, обрабатывает их и передает на один из следующих слоев. Выходной слой дает окончательный результат обработки всех данных искусственной нейронной сетью.

Архитектура нейросетей, включающая скрытые слои, – главное преимущество нейронных сетей перед классическими видами машинного обучения. Благодаря ей нейросети **точнее моделируют взаимосвязи между входными и выходными данными и решают более сложные задачи** на основе менее структурированных данных, таких как текст, звуковые сигналы и изображения. Так, по оценкам экспертов, алгоритмы на основе глубокого обучения выполняют задачи по классификации изображений на 41% точнее алгоритмов традиционного машинного обучения, по распознаванию лиц – на 27%, по распознаванию голоса – на 25%¹⁹.

Вместе с тем архитектура нейронных сетей определяет и ключевые ограничения их использования. Наличие большого количества скрытых слоев предъявляет более высокие требования к вычислительным ресурсам и объемам данных. Как следствие, обучение, тестирование и работа нейросетей требуют более значительных временных затрат и инвестиций. Кроме того, одной из основных проблем использования является проблема объяснимости алгоритмов, лежащих в основе нейронных сетей (так называемая проблема черного ящика).

Ключевые функциональные направления использования глубокого обучения:

- **Компьютерное зрение (Computer Vision)** – технология, направленная на извлечение информации из визуальных данных и принятие решений на основе этой информации. Технология используется для распознавания лиц и жестов, поиска и извлечения информации на изображениях, интеллектуальной обработки документов. В финансовом секторе компьютерное зрение применяется, например, для *распознавания информации на бумажных носителях и перевода ее в цифровой вид*.
- **Обработка естественного языка (Natural Language Processing)** – технология ИИ, которая позволяет компьютерам понимать (Natural language understanding) и генерировать (Natural language generating) текстовую информацию на естественных языках. В числе ключевых областей применения – *чат-боты, системы машинного перевода, генерация текстового контента и прочее*. Генеративный ИИ также использует технологию обработки естественного языка. Так, вне зависимости от создаваемого таким ИИ контента (текстового или визуального), если пользователь для получения контента должен вводить текстовый запрос, такая модель использует обработку естественного языка.
- **Распознавание и синтез речи (Automatic Speech Recognition & Text-To-Speech)** – комплекс технологий, направленный на преобразование речи человека в цифровую информацию и имитацию речи. На финансовом рынке данные технологии применяются для разработки *голосовых чат-ботов, виртуальных помощников, автоматизации обзвонки клиентов для информирования о задолженности (роботы-коллекторы), речевой аналитики, аутентификации клиентов*²⁰.

¹⁹ [An executive's guide to AI](#). McKinsey (2022).

²⁰ [Голос в помощь](#). Финансовая сфера. Банковское обозрение (Июль 2022).

Многие эксперты в качестве отдельного функционального направления использования ИИ в рамках описанных выше технологий выделяют **роботизированную автоматизацию бизнес-процессов (Robotic Process Automation, RPA)** – комплекс решений, основанный на использовании программных роботов, технологий компьютерного зрения и обработки естественного языка, нацеленных на автоматизацию рутинных задач, таких как *извлечение данных из документов, перенос данных из одной системы в другую, выставление счетов и так далее*. Глобальный рынок роботизации бизнес-процессов динамично растет. По прогнозам экспертов, средний годовой темп его прироста в ближайшие годы составит 39,9%²¹.

По оценкам экспертов, мировой финансовый сектор находится в числе отраслей, которые наиболее активно внедряют технологии ИИ. Так, технологии ИИ для роботизации бизнес-процессов используют 47% финансовых компаний, обработки естественного языка – 42%, распознавания речи – 30%²².

ДОЛЯ КОМПАНИЙ, ИСПОЛЬЗУЮЩИХ ТЕХНОЛОГИИ ИИ В РАЗНЫХ ОТРАСЛЯХ*
(%)

Табл. 1

| | Компьютерное зрение | Глубокое обучение | Распознавание лиц | Генерация текста | Распознавание речи | Обработка естественного языка | Рекомендательные системы | Обучение с подкреплением | Роботизация бизнес-процессов |
|------------------------------------|---------------------|-------------------|-------------------|------------------|--------------------|-------------------------------|--------------------------|--------------------------|------------------------------|
| Все индустрии | 34 | 30 | 18 | 18 | 23 | 33 | 25 | 20 | 39 |
| Деловые и юридические консультации | 32 | 37 | 11 | 12 | 22 | 34 | 23 | 26 | 46 |
| Розничная торговля | 33 | 36 | 19 | 20 | 11 | 22 | 32 | 19 | 25 |
| Финансовые сервисы | 24 | 22 | 24 | 20 | 30 | 42 | 30 | 19 | 47 |
| Здравоохранение | 32 | 18 | 5 | 5 | 12 | 29 | 16 | 13 | 16 |
| Хайтек, телеком | 37 | 45 | 16 | 24 | 29 | 40 | 34 | 23 | 48 |

* [Artificial Intelligence Index Report](#). Stanford University (2023).

Развитие перспективных методов машинного обучения

С точки зрения развития технологии машинного обучения эксперты²³ выделяют в том числе следующие направления и их комбинации:

- **Федеративное обучение (Federated Learning)** – способ обучения моделей ИИ на основе нескольких независимых наборов данных. При федеративном обучении используется децентрализованный подход, при котором ИИ обучается не на одном общем пуле данных, а на данных из разных источников. Развитие федеративного обучения позволит компаниям, использующим ИИ, тренировать свою модель, не собирая персональные и иные конфиденциальные данные с устройств (или из иных источников) пользователей. На финансовом рынке федеративное обучение может быть использовано в кредитном скоринге заемщиков, применении рекомендательных систем²⁴, при противодействии мошенничеству, при аналитике данных колл-центров финансовых организаций, а также в других традиционных моделях ИИ.

²¹ [Robotic Process Automation Market Size, Share & Trends Analysis Report by Type, by Application, by Development, by Organization, by Region, and Segment Forecast, 2023–2030](#). Grand View Research (2023).

²² [Artificial Intelligence Index Report](#). Stanford University (2023).

²³ [Применение технологий искусственного интеллекта на финансовом рынке](#). Ассоциация ФинТех (Октябрь 2023); [What's New in the 2023 Gartner Hype Cycle for Emerging Technologies](#) (August 2023).

²⁴ Рекомендательные системы – сервисы, предоставляющие рекомендации пользователю по продуктам, услугам, контенту на основании данных о его профиле, а также информации о предшествующей пользовательской активности.

- **Мультимодальное обучение (Multimodal Learning).** Посредством мультимодального глубокого обучения генеративные модели ИИ могут предоставлять ответ с использованием различных форматов информации – визуальной, аудио, текстовой. Алгоритмы на основе такого вида обучения, могут, например, на текстовый запрос дать ответ не только в виде текста, но и в виде изображения. Примеры подобных решений уже существуют на российском рынке.
- **Обучение Low-code/No-code** – подход к разработке программного обеспечения, который позволяет создавать приложения либо с минимальным написанием кода, либо без него. Подход Low-Code в ИИ предоставляет готовые инструменты и компоненты, которые можно легко настроить и объединить при создании ИИ-приложений. Вместо классического подхода к разработке в рамках Low-code приложения могут создаваться с использованием графического интерфейса и визуальных инструментов. По мере развития использования Low-code/No-code разработка традиционных моделей ИИ становится доступнее для участников рынка, поскольку требует найма менее высокооплачиваемых специалистов.

2. ГЛОБАЛЬНЫЕ ТРЕНДЫ РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Общие тренды развития искусственного интеллекта

По прогнозам экспертов, мировой рынок ИИ-решений в 2023–2030 годах вырастет с 196,6 млрд до 1,8 трлн долларов США со среднегодовым темпом роста 37%²⁵. Согласно оценке экспертов McKinsey, общий экономический потенциал внедрения ИИ в разных отраслях составляет до 17,1–25,6 трлн долларов США. Прирост будет обеспечиваться за счет увеличения продуктивности бизнеса благодаря автоматизации рабочих процессов, роста производительности труда и повышения спроса на продукты с технологией ИИ в связи с ростом качества этих продуктов²⁶.

Такой эффект будет возможен в том числе за счет снижения стоимости внедрения в деятельность организаций моделей ИИ, а также повышения доступности данных для обучения таких моделей.

К общим трендам относятся²⁷:

- **Демократизация технологий ИИ** – удешевление, повышение распространенности и доступности ИИ-решений для широкого круга пользователей и организаций, в том числе предприятий малого и среднего бизнеса. Этот процесс происходит благодаря распространению сервисов обучения Low-code/No-code, описанных выше, а также развитию свободного программного обеспечения, повышению доступности облачных сервисов.
- **Конвергенция IoT²⁸ и ИИ** – синергия двух указанных технологий, позволяющая smart-устройствам (включая мобильные телефоны) собирать данные, которые затем ИИ может использовать для обучения. По мере общего роста количества smart-устройств увеличивается и объем данных, которые смогут использовать ИИ.

Также внедрению ИИ может способствовать повышение доверия к нему со стороны пользователей за счет роста популярности концепции ответственного искусственного интеллекта (Responsible AI), которая включает в себя комплекс технологических, управленческих и операционных решений, направленных на минимизацию предвзятости, увеличение объяснимости, безопасности, конфиденциальности и соблюдение нормативных требований при разработке и использовании ИИ.

Эксперты отдельно выделяют распространение принципов AI TRiSM (AI trust, risk and security management)²⁹ в качестве одного из ключевых технологических трендов, связанных с ИИ. Данная концепция предполагает создание надежных моделей, обеспечивающих справедливость по отношению к людям различных групп, а также надежность, эффективность и безопасность данных, что обеспечивает доверие конечного потребителя к технологии и, как следствие, способствует увеличению ее распространения.

Развитие генеративного искусственного интеллекта

Ключевой технологический тренд, связанный с использованием ИИ, – развитие и широкое распространение генеративного ИИ, в том числе больших языковых моделей. Генеративный ИИ используется для создания контента, включая текст, аудио, изображения, видео.

²⁵ [Artificial Intelligence Market to Hit \\$ 1,811.75 Billion by 2030: Grand View Research](#). Bloomberg (July 2023).

²⁶ [The economic potential of generative AI: The next productivity frontier](#). McKinsey (June 2023).

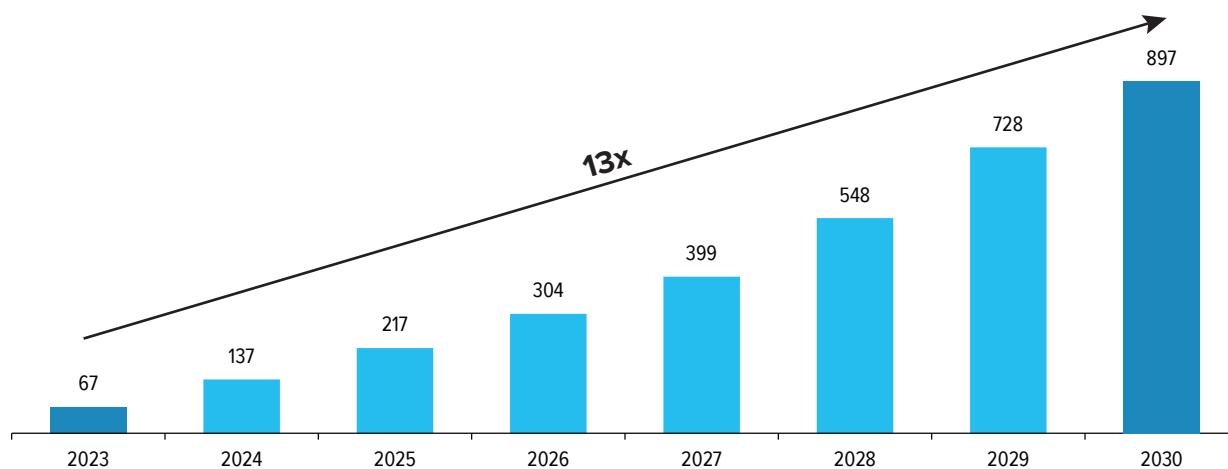
²⁷ [Применение технологий искусственного интеллекта на финансовом рынке](#). Ассоциация ФинТех (Октябрь 2023).

²⁸ IoT (Internet of things) – множество физических объектов, подключенных к Интернету и обменивающихся данными друг с другом.

²⁹ [Gartner – Top Strategic Technology Trends 2023](#). Gartner (October 2022). AI trust, risk and security management – управление безопасностью, рисками и доверием к искусственному интеллекту.

ОЦЕНКА ОБЪЕМА РЫНКА ГЕНЕРАТИВНОГО ИИ*
(МЛРД ДОЛЛ. США)

Рис. 1



* Generative AI to Become a \$ 1.3 Trillion Market by 2032. Research Finds. Bloomberg (June 2023).

По мнению экспертов, **в ближайшие годы генеративный ИИ станет одним из самых динамично растущих направлений ИИ**. По оценке Bloomberg Intelligence, глобальный рынок решений на основе генеративного ИИ в 2023–2030 годах увеличится более чем в 13 раз – с 67 млрд до 897 млрд долларов США³⁰. Рост публичного интереса к сфере ИИ в последние годы во многом связан с развитием моделей генеративного ИИ.

Благодаря стремительному росту производительности модели генеративного ИИ имеют широкий потенциал применения в различных сферах в части оптимизации процесса взаимодействия с пользователями, улучшения клиентского опыта и других.

По оценкам McKinsey, экономический эффект от внедрения продуктов генеративного ИИ в разных секторах мировой экономики составит 2,6–4,4 трлн долларов США в год, а 75% прироста даст использование генеративного ИИ в маркетинге, клиентских сервисах и разработке программного обеспечения. **Финансовый сектор – в числе отраслей, которые получают наибольшую экономическую выгоду от внедрения генеративного ИИ**, который позволит увеличить выручку банков на 200–340 млрд долларов США в год, а страховых компаний – на 50–70 млрд долларов США в год. **По мнению экспертов, наиболее перспективные направления применения генеративного ИИ в финансовом секторе – развитие клиентских сервисов (чат-боты и голосовые помощники), управление рисками и комплаенс**³¹.

В качестве отдельного направления развития генеративного ИИ выделяются **большие языковые модели** (Large Language Model, LLM) – системы генерации текстовой информации на основе нейросетей архитектуры трансформер. Эти нейросети обрабатывают последовательности неструктурированных текстовых данных при помощи механизма внимания (Attention Mechanism), что позволяет модели выделять нужную информацию из истории обработки данных и учитывать ее при обработке новых данных. Примеры таких моделей есть как в России (Сбер – GigaChat и «Яндекс» – YandexGPT), так и в других странах (США – ChatGPT, Китай – ERNIE Bot).

Для обучения больших языковых моделей требуются значительные объемы текстовой информации, доступной в Интернете. Например, по оценкам отраслевых экспертов, GPT-3³² был обучен примерно на 45 терабайтах текстовых данных, что эквивалентно пример-

³⁰ [Generative AI to Become a \\$ 1.3 Trillion Market by 2032](#). Bloomberg (June 2023).

³¹ [The economic potential of generative AI: The next productivity frontier](#). McKinsey (June 2023).

³² GPT-3 (Generative Pre-trained Transformer 3) – третье поколение большой языковой модели от OpenAI.

но 300 млн страниц текстовой информации в стандартном цифровом формате³³. Обработка столь значительного объема данных требует вычислительных мощностей и существенных инвестиций в разработку таких моделей.

Одно из перспективных решений – **генерация синтетических данных**, то есть искусственных данных, которые не имеют отношения к конкретному клиенту или пользователю, созданы специально обученной моделью ИИ, имитируют характеристики реальных клиентских данных 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. По оценкам Gartner³⁴, к 2030 году синтетические данные могут полностью заменить потребность в клиентских данных для обучения моделей ИИ.

Эксперты также отмечают, что развитие генеративного ИИ приведет к росту количества **дипфейков (Deepfakes)**, которые представляют собой технологии машинного синтеза изображений, аудио- и видеозаписей с целью создания видеоизмененного, но реалистичного контента. К числу негативных последствий распространения дипфейков относят увеличение рисков манипулирования информацией, распространение дезинформации. В рамках зарубежных законодательств разрабатываются механизмы правового регулирования применения дипфейков.

³³ [How Big Will AI Models Get?](#). Institute of Electrical and Electronics Engineers (February 2023).

³⁴ [Top Strategic Technology Trends 2023](#). Gartner, Inc. (June 2023).

3. ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА ФИНАНСОВОМ РЫНКЕ

До недавнего времени на рынке применялись преимущественно традиционные решения в области ИИ. Однако за последнее время возможности компьютерной обработки данных и вычислительные мощности значительно выросли, что создало условия и для активного развития и применения генеративного ИИ.

В настоящее время наибольшее распространение на финансовом рынке получили решения с использованием традиционного ИИ. Такие модели были усовершенствованы настолько, что в некоторых случаях могут применяться без привлечения узкоспециализированных экспертов. ИИ постепенно становится технологией общего назначения и способен трансформировать бизнес-процессы и рыночные практики финансовых организаций (табл. 2).

ПРИМЕРЫ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ФИНАНСОВЫХ ОРГАНИЗАЦИЯХ

Табл. 2

| Область применения | Сфера применения | Тип ИИ |
|---|--|---|
| Первая линия: фронт-офис | Чат-боты | ИИ-алгоритмы на основе обработки естественного языка |
| | «Умные» инструменты маркетинга | ИИ-алгоритмы с элементами глубокого обучения для формирования персональных предложений |
| | Алготрейдинг, инвестиционное консультирование и оценка стоимости активов | ИИ-алгоритмы на основе регрессионных моделей для оценки показателей |
| Вторая линия: проверка и обработка операций | Скоринг | Приложения на основе ИИ-алгоритмов для проведения кредитного скоринга, доступные в том числе клиентам |
| | Подтверждение операций, обработка документов | ИИ-алгоритмы, подключенные к сведениям о платежных данных и системе управления рисками для мгновенного принятия решений |
| Третья линия: операционный учет | Мониторинг транзакций | ИИ-алгоритмы, способные в режиме реального времени отслеживать подозрительные операции и оповещать пользователей |
| Общее | Разработка, оптимизация ПО | Сквозная сервисная функция для бизнес-процессов |

Важно отметить, что развитие инструментов ИИ на финансовом рынке в значительной степени происходит в русле общей траектории совершенствования методов применения ИИ, как и иных методов и алгоритмов. В то же время существует ряд специфических аспектов применения технологии ИИ для финансовой сферы.

Масштабное применение ИИ может потенциально рассматриваться в качестве дополнительного источника повышения эффективности как для финансовых организаций, так и для их клиентов.

Важным эффектом применения ИИ являются сокращение затрат при масштабировании операционной деятельности, оптимизация ресурсов, повышение лояльности клиентов и доходности продаж за счет персонализированных продуктов. Расширенная аналитика с встроенными алгоритмами машинного обучения может помочь совершенствовать модели предложения продуктов и их обслуживания.

Вместе с тем расширение использования традиционного ИИ, а также активное развитие генеративного ИИ может привести к трансформации бизнес-моделей финансовых организаций.

При этом применение ИИ может рассматриваться как дополнительный источник уязвимости для финансового рынка, поскольку повышается возможность сбоев и искажений бизнес-процесса от ожидаемого результата за счет специфики вариативности решений инструментов ИИ, особенно в периоды стрессов, что в итоге может привести к ухудшению клиентского опыта и ущербу для потребителей. Этим объясняется внимание и осторожное отношение регуляторов к применению технологий ИИ на финансовом рынке.

Помимо влияния на клиентский опыт и ущерба потребителям, применение ИИ, особенно в случаях его использования финансовыми организациями в анализе данных, при углубленном скоринге может привести к реализации операционного риска и, соответственно, непредвиденным убыткам финансовых организаций.

Использование искусственного интеллекта в России

Согласно результатам опроса, проведенного Ассоциацией ФинТех¹, в настоящее время в бизнес-процессах 95% финансовых и технологических организаций – респондентов применяются системы на основе ИИ. По данным опроса, 87% респондентов уже используют ИИ-решения для анализа данных, 63% компаний – для работы с текстом, 35% компаний используют речевые технологии, 30% компаний применяют решения ИИ в области компьютерного зрения.

¹ [Применение технологий искусственного интеллекта на финансовом рынке](#). Ассоциация ФинТех (Октябрь 2023).

Одним из основных направлений использования ИИ на финансовом рынке является **ско-ринг**, который позволяет финансовым организациям более точно сегментировать клиентов на однородные группы, определять их потребности и риск-профиль. Системы оценки кредитоспособности, управляемые с помощью ИИ, могут быть полностью автоматизированы и предлагать моментальные решения на основе проанализированных данных, дополняя традиционные методы оценки. Например, анализ оплаты счетов за коммунальные услуги позволяет добавить в модель оценки кредитоспособности новые переменные и точнее рассчитать кредитный рейтинг для клиентов, у которых недостаточно кредитной истории.

Появление искусственного интеллекта на рынке

Одним из первых успешных приложений с применением ИИ, получивших широкое распространение, стала система автоматической авторизации кредитов, внедренная розничным продавцом Marks & Spencer в 1991 году в Великобритании. Программа Behavioural Scoring позволяла автоматически обрабатывать около 90% кредитных заявок и аналогичных операций и предлагать решение на основе анализа данных о месте жительства, кредитной истории и операциях по счету¹.

¹ [Artificial Intelligence market and capital flows: Artificial Intelligence and the financial sector at crossroads](#). European Parliament (May 2021).

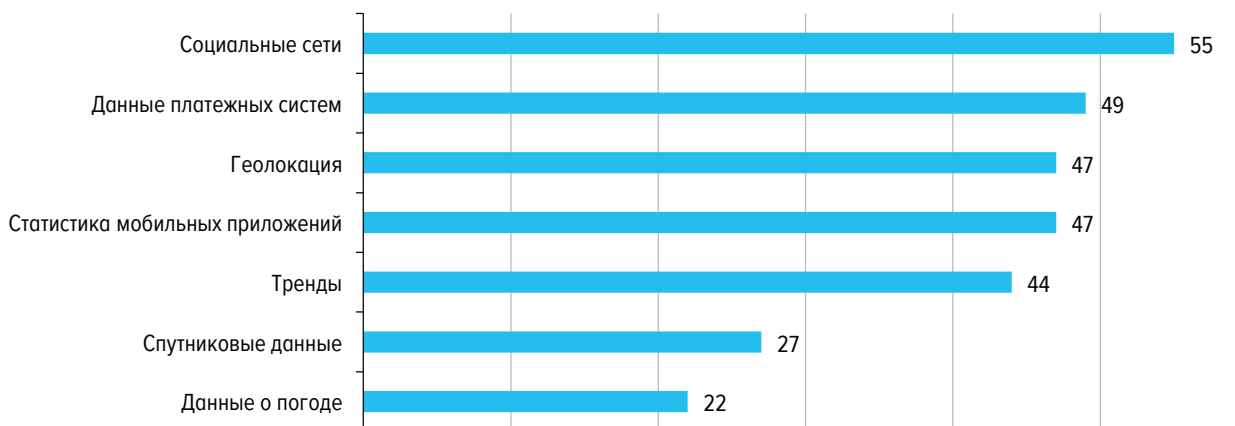
Такой подход формирует условия для создания более комфортной экосистемы для клиентов, где им доступны не только отдельные персонализированные продукты, но и комплексный финансовый сервис, превентивно ориентированный на их потребности.

«Умный» скоринг на основе ИИ позволяет эффективно использовать не только финансовую информацию, такую как кредитная история или располагаемый доход, но и иные, так называемые альтернативные данные о клиенте: финансовые привычки, геолокация, сведения из социальных сетей и так далее (рис. 2).

Использование «альтернативных» данных, равно как и само применение моделей ИИ, может быть полезным при поиске новых взаимосвязей или скрытых зависимостей. Вместе с тем применение таких данных в финансовом секторе вызывает ряд опасений как у клиентов,

ПРИМЕНЕНИЕ «АЛЬТЕРНАТИВНЫХ» ДАННЫХ ДЛЯ СКОРИНГА*
(% ОТ КОЛИЧЕСТВА ОПРОШЕННЫХ ОРГАНИЗАЦИЙ)

Рис. 2



* *Transforming Paradigms. A Global AI in Financial Services Survey*. World Economic Forum. Cambridge Centre for Alternative Finance (January 2020).

так и у регуляторов в части защиты конфиденциальности и этичности использования. По этой причине законодательство некоторых стран ограничивает неконтролируемое применение в алгоритмах ИИ таких сведений, опасаясь непреднамеренной дискриминации заявителей или иного ущерба для них³⁵.

Противодействие мошенничеству – еще одно направление, в котором расширяется использование механизмов ИИ. Благодаря способности анализировать большие объемы данных, выявлять нетипичные транзакции и отслеживать аномалии, труднозаметные для человека, технологии ИИ позволили за счет снижения ресурсоемкости вывести процедуры обнаружения фактов мошенничества на новый уровень, превосходящий классический анализ контрольных факторов рисков.

В режиме реального времени алгоритмы ИИ способны маркировать события как подозрительные и блокировать нехарактерные для клиента транзакции (например, с признаками мошенничества с кредитными картами) или подключать сотрудника финансовой организации для проведения более детальной проверки и анализа документов и изображений (страховое мошенничество, искажения в финансовой отчетности, процедуры KYC³⁶).

Вместе с тем важно отметить, что мошенничество – это нетипичная транзакция, которая не поддается непосредственному наблюдению. Резкое изменение поведения клиента может привести к нарушению привычных взаимосвязей, сбоям в работе алгоритма, а значит, к прямо противоположному результату. По этим причинам многие финансовые организации по-прежнему полагаются на традиционные системы проверки и контроля или совмещают обе опции.

Развитие ИИ потенциально может значительно изменить подходы RegTech³⁷ при **управлении рисками** в финансовых организациях.

Результативность алгоритмов тестируется при управлении рисками ПОД/ФТ и кибербезопасности³⁸. В частности, в разрезе кибербезопасности ИИ может помочь настроить системы безопасности организации, которые смогут идентифицировать аномальное поведение внутренних систем и вредоносные программы и мгновенно реагировать на них.

³⁵ [General Data Protection Regulation](#). Regulation EU 2016/679 of the European Parliament and of the Council of 27 April 2016.

³⁶ KYC (Know Your Customer) – термин, означающий необходимость идентифицировать и установить личность контрагента перед проведением финансовой операции.

³⁷ RegTech (Regulatory Technology) – технологии, используемые для упрощения выполнения требований регуляторов.

³⁸ [Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance](#). IMF (October 2021).

СФЕРЫ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ИНВЕСТИЦИОННОМ СЕКТОРЕ*

Табл. 3

| Торговля и управление активами | Инфраструктура инвестиционного сектора (торговая, учетная, расчетно-клиринговая) |
|---|--|
| <ul style="list-style-type: none"> • Консультирование и сопровождение клиентов (в том числе инвестиционные рекомендации, подбор инвестиционных продуктов и инструментов, формирование торговых стратегий) • Алгоритмическая торговля • Управление активами и ребалансировка инвестиционного портфеля • Управление рисками (индикаторы раннего предупреждения дефолтов, оценка риск-профиля, анализ волатильности на рынке, оценка риска ликвидности, оценка качества исполнения сделок) | <ul style="list-style-type: none"> • Управление рисками (идентификация и анализ аномальных активностей в торговле, оценка рисков неисполнения сделок, прогнозирование ликвидности) • Операционная деятельность (оценка качества данных, подтверждение данных, автоматическая обработка неструктурированных данных) |

* Составлено с учетом материалов: [TRV Risk Analysis Artificial intelligence in EU securities markets](#). ESMA (February 2023); [Unleashing the potential of AI in securities services](#). Deutsche Bank (June 2021).

Развитие ИИ может также *повлиять на управление инвестиционными и кредитными рисками*, предоставив дополнительные инструменты для их мониторинга и оценки. В режиме реального времени ИИ-алгоритмы способны отслеживать множество факторов риска и тестировать эффективность и устойчивость портфеля к различным рыночным/экономическим сценариям. В частности, они могут сформировать систему индикаторов раннего предупреждения о потенциальном неисполнении клиентами обязательств (дефолте) и динамические показатели измерения риск-профиля клиента.

Некоторые маркетмейкеры внедряют модели ИИ для минимизации риска обесценения свободных финансовых активов (*Inventory Risk*) и максимизации рентабельности своего баланса. Отдельные управляющие активами и хедж-фонды автоматизируют процессы управления рисками и обеспечения соблюдения регуляторных требований (комплаенс) путем отслеживания поведения управляющих индивидуальными портфелями активов, автоматизации ежедневных отчетов о качестве исполнения сделок (*Execution Quality Reports*) и процедуры оценки риска рыночной ликвидности.

Технологии ИИ получают все большее распространение среди институциональных инвесторов, брокеров, доверительных управляющих и инвестиционных советников. Основные направления проникновения механизмов ИИ в инвестиционной области приведены в таблице 3. Отмечаются также практики применения технологии ИИ для проведения оценки активов, в том числе для определения готовности непубличных компаний к IPO и анализа инвестиционной привлекательности объектов недвижимости³⁹.

При управлении инвестиционным портфелем алгоритмы ИИ способны быстрее спрогнозировать доходность инструментов, чем традиционные. Это позволяет оперативно сформировать второе мнение, подготовить рекомендации при формировании и ребалансировке портфеля, автоматизировать технический и фундаментальный анализ – например, извлекать экономически значимую информацию из корпоративной отчетности (методы обработки естественного языка) или построить регрессионную модель для оценки индикаторов как на общеэкономическом уровне, так и для отдельных компаний. Сложным алгоритмам под силу учитывать изменения веса и роли совокупности факторов на различных этапах анализа (природные факторы, снижение/увеличение темпов роста ведущих экономик мира и так далее), однако они могут ошибаться в прогнозе при возникновении непредвиденных шоков. По этой причине финансовые организации предпочитают применять ИИ параллель-

³⁹ Составлено с учетом материалов: [TRV Risk Analysis Artificial intelligence in EU securities markets](#). ESMA (February 2023); [Unleashing the potential of AI in securities services](#). Deutsche Bank (June 2021).

но с классическими методами для сравнения и оценки полученных результатов (например, при подготовке рейтинга или матрицы рисков) на этапах пред- и постторгового анализа.

При ребалансировке инвестиционного портфеля может применяться так называемый эволюционный алгоритм (Evolutionary Algorithm), который позволяет решать задачи оптимального распределения активов при их небольшом количестве в портфеле. Он учитывает в оценке переменную модельного риска (в частности, рассчитывает коэффициенты волатильности для каждого актива в результате ошибочной спецификации модели), что значительно снижает ошибки в прогнозе. Портфели, для которых применяется такая методика, как правило, показывают лучший коэффициент Шарпа⁴⁰ (примерно на 10%), чем те, которые не учитывают модельные риски⁴¹.

Применение ИИ в **алгоритмической торговле** стало возможным с развитием вычислительных технологий и методов МО. Компьютеры стали способны отбирать и анализировать большие объемы данных, отслеживать изменения на рынке в реальном времени и совершать сделки значительно быстрее и зачастую эффективнее человека.

ИИ может использоваться для построения моделей, которые будут либо помогать человеку в процессе принятия решений (например, посредством определенных предупреждающих сигналов), либо автономно выполнять операции с учетом конкретного сценария и набора обстоятельств. Как правило, они основаны на техническом анализе и могут прогнозировать рыночную динамику и тренды, а также рассчитывать транзакционные издержки по сделкам.

Несмотря на преимущества скорости оценки рыночных показателей, торговля с применением ИИ-алгоритмов имеет ряд недостатков. Так, например, большая чувствительность ИИ-моделей к сильным рыночным шокам может создавать эффект автокорреляции алгоритмических систем и усиливать системные риски⁴².

Риски Flash Crash

Например, таким системным риском стал торговый эпизод 6 мая 2010 года (*Flash Crash*), когда при помощи автоматического трейдинга с применением стратегии «имитация заявок» (*spoofing*) были размещены и затем сразу отменены тысячи крупных ордеров на контракты на индекс S&P 500. Автоматические заявки повлияли на биржевые котировки, которые учитывались трейдерами и алгоритмами, и привели к падению рыночных индексов, остановке торгов и изменению капитализации рынка на 1 трлн долл. США.

В целом в открытых источниках отсутствует упоминание об эффективности алгоритмических роботов на длительном горизонте инвестирования – такой инструментарий, как правило, используется для краткосрочных торговых стратегий в качестве «умного» и быстрого калькулятора. В вопросе же использования обучаемых ИИ-алгоритмов для долгосрочных инвестиций может возникнуть ряд сложностей, связанных со сложностью использования исторических рядов цен в шоковые периоды для обучения торговых алгоритмов.

Роботы-советники являются естественной нишей масштабирования торговых алгоритмов для развития и применения технологий ИИ. Такие автоматизированные системы предлагают инвестиционные рекомендации и стратегии, адаптированные к конкретным обстоятельствам и индивидуальным предпочтениям клиентов финансовых институтов в зависимости от их инвестиционных целей и уровня принимаемого риска.

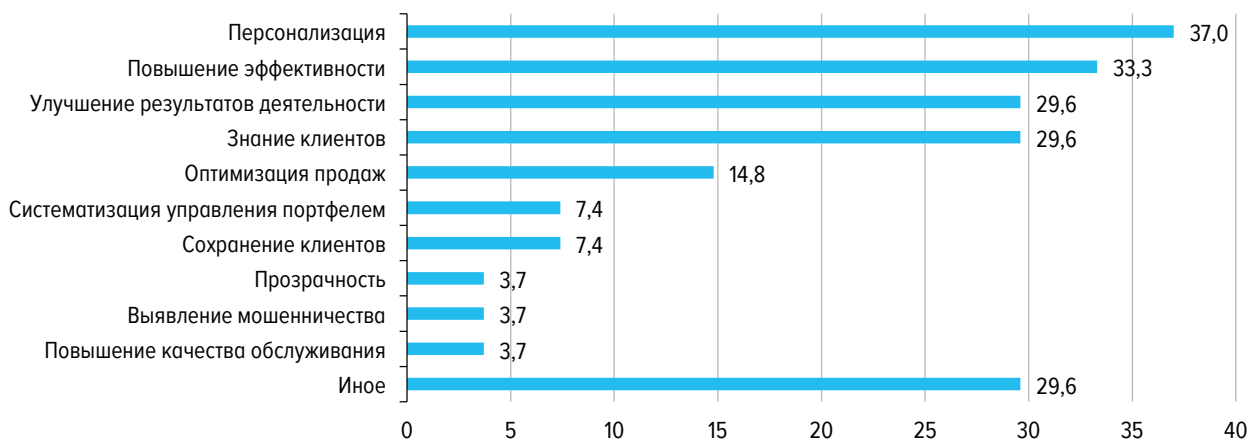
⁴⁰ Коэффициент Шарпа – показатель эффективности инвестиционного портфеля (актива), который вычисляется как отношение средней премии за риск к среднему отклонению портфеля.

⁴¹ [Artificial Intelligence in Asset Management](#). CFA. 2020.

⁴² [The Flash Crash: High-Frequency Trading in an Electronic Market](#). A. Kirilenko, A. Kyle, M. Samadi *Journal of Finance* 72 (3).

АКТУАЛЬНЫЕ НАПРАВЛЕНИЯ, СВЯЗАННЫЕ С ВНЕДРЕНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА*
(% КОМПАНИЙ ИЗ ЧИСЛА РЕСПОНДЕНТОВ, ПОДТВЕРДИВШИХ АКТУАЛЬНОСТЬ)

Рис. 3



* На основе анкетирования компаний Германии, применяющих роботов-советников. *How much Artificial Intelligence do Robo-Advisers really use?* C. Their, D. dos Santos Monteiro. 2022.

В соответствии с исследованиями основными трендами в интеграции ИИ в такие программы являются персонализация продуктов и услуг (37%), повышение эффективности (33,3%) и результатов деятельности (29,6%) и знаний о клиентах (29,6%) (рис. 3).

Большинство роботов-советников используют простые, основанные на правилах (дедуктивные) алгоритмы, однако некоторые начинают применять предиктивные алгоритмы. Работа ИИ не исключает вмешательства человека (Manual Intervention) в процесс предоставления торговой рекомендации – в зависимости от заранее настроенных параметров, например повышенной волатильности рынка или объема транзакции, система может быть переведена в ручной режим.

Сложности возникают, если в программу интегрирована автоматическая функция заключения сделок, поскольку это может привести к чрезмерной торговле и дополнительным комиссиям. Как правило, роботы-советники также не учитывают налоги и инфляцию. Следовательно, необходимо осмотрительно выбирать такую программу и обращать внимание на технические характеристики и ограничения моделей, в частности на раскрытие информации о применении технологий ИИ и машинного обучения в конкретных моделях.

Вместе с тем в инфраструктурных решениях ИИ используется в ограниченном периметре операций⁴³, например для автоматического разбора неструктурированных сообщений при обработке корпоративных действий, для отслеживания проблем в функционировании ИТ-решений⁴⁴, а также для прогнозирования нагрузки на ИТ-системы, раннего предупреждения рисков сбоев и иных инцидентов.

⁴³ По информации Европейского управления по надзору за рынком ценных бумаг и финансовыми рынками (ESMA), только единичные депозитарии и биржи применяют ИИ в своей деятельности. Подробнее см. ESMA. [TRV Risk Analysis Artificial intelligence in EU securities markets](#) (February 2023).

⁴⁴ [Искусственный интеллект в национальном депозитарии](#). НКО АО НРД.

Искусственный интеллект как фактор развития финансового рынка

В перспективе развитие ИИ может привести к трансформации бизнес-процессов и бизнес-моделей участников финансового рынка.

Одним из таких примеров может стать страховой рынок. Беспилотные автомобили могут менять характер ответственности в автостраховании. Кто несет ответственность в случае ДТП: пассажир, производитель автомобиля или разработчик программного обеспечения алгоритмов ИИ? Такая постановка вопроса может привести к трансформации традиционного автострахования в том виде, в котором оно есть сейчас.

Использование беспилотных транспортных средств предполагает сокращение количества аварий. Однако выход из строя самой системы ИИ или хакерская атака может вызвать каскадную серию происшествий, приводящую к значительному увеличению максимально возможного ущерба, что в целом формирует переход от высокочастотных небольших рисков к низкочастотным, но существенным.

Такое накопление рисков представляет собой серьезное препятствие для страхования киберрисков. Риски высокой степени тяжести требуют очень высоких лимитов покрытия и выплат, что ставит перед страховыми компаниями задачу пересмотреть традиционное страховое покрытие и разработать инновационные страховые продукты.

Совокупность непрерывного андеррайтинга и поступающие данные от телематических или носимых устройств позволят создать страхование по требованию для различных активов. Страхование покрытие можно будет приобрести для широкого спектра рисков низкой степени тяжести на время фактического использования актива (например, дополнительное страхование личных вещей от кражи или повреждения, новые продукты в туристическом или автостраховании), которые раньше не поддавались страхованию.

ИИ, внедряемый в других сферах экономики, приводит к размыванию традиционных отраслевых границ и созданию глобальных экосистем. Такие экосистемы позволят страховым компаниям не только выйти на новые потоки доходов, пересмотрев свою традиционную роль в экономике, но и интегрировать свои страховые продукты в повседневную жизнь клиента, удовлетворяя его разнообразные потребности в рамках единого взаимодействия.

В сфере мониторинга **обращений клиентов и урегулирования претензий ИИ**⁴⁵ может применяться для классификации поступающих обращений, автоматического считывания и обработки данных. Развитие генеративного ИИ существенно упрощает и оптимизирует взаимодействие финансовой организации и заявителя. Чат-боты и голосовые помощники принимают на себя рутинные функции на протяжении всего процесса урегулирования претензии.

Широкое распространение такие практики могут получить в страховом секторе. Алгоритмы могут применяться для классификации поступающих обращений, автоматического считывания и обработки данных, например для извлечения информации из медицинских карт, распознавания типов транспортных средств и повреждений. В целом при сборе и анализе больших объемов исторических данных ИИ может применяться:

- для сегментации претензий по виду и типу сложности;
- для определения уровня компенсации при наступлении страховых случаев;
- для анализа причин отклонения претензий и/или изменения суммы страхового возмещения.

Потенциально ИИ может быть интегрирован в систему обработки заявок страховой компании и автоматически принимать решение в простых случаях.

⁴⁵ Составлено с учетом материалов: [Transforming Paradigms. A Global AI in Financial Services Survey. World Economic Forum. Cambridge Centre for Alternative Finance. \(January 2020\)](#); [The Impact of big Data and Artificial Intelligence \(AI\) in the Insurance Sector. OECD \(2020\)](#).

4. УСЛОВИЯ РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА РОССИЙСКОМ ФИНАНСОВОМ РЫНКЕ

Внедрение технологий ИИ в ключевых отраслях экономики ведет к значимому росту эффективности деятельности организаций. По оценкам экспертов, внедрение ИИ в различных отраслях российской экономики может дать дополнительный прирост ВВП страны на 2% в 2025 году⁴⁶. Финансовый рынок наиболее активен во внедрении таких технологий. Согласно данным исследования Ассоциации ФинТех, технологии ИИ используют в своей деятельности более 95% российских финансовых компаний – участников исследования⁴⁷.

Дальнейшее развитие и внедрение ИИ в деятельность финансовых организаций зависит в том числе от создания таких условий, как:

- 1) повышение доступности и качества данных, необходимых для обучения и тестирования моделей ИИ, включая:
 - доступ участников финансового рынка к государственным данным и развитие Цифрового профиля;
 - развитие механизмов обработки персональных данных и других сведений, относящихся к различным видам тайн;
 - развитие применения технологии открытых программных интерфейсов для передачи данных (Открытые API), создание Платформы коммерческих согласий;
- 2) развитие и повышение доступности программного обеспечения и вычислительной инфраструктуры, необходимой для решения задач в области ИИ, включая:
 - формирование высокопроизводительной вычислительной инфраструктуры для разработки решений в области ИИ;
 - использование программного обеспечения, распространяемого на основе открытых лицензий;
 - развитие отечественного программного обеспечения, используемого для обучения ИИ;
- 3) развитие научной базы в области ИИ;
- 4) развитие профессиональных компетенций в области ИИ;
- 5) обеспечение доверия граждан к технологиям ИИ;
- 6) развитие конкуренции через широкое применение технологии ИИ на рынке, в том числе небольшими компаниями;
- 7) создание регуляторных условий, способствующих развитию применения технологии с учетом возникающих рисков.

Повышение доступности и качества данных

Доступ участников финансового рынка к государственным базам данных

Помимо сведений от внутренних источников, финансовые организации заинтересованы в получении различных внешних данных, в частности от государственных органов. По информации Банка России, финансовые организации предъявляют наибольший спрос на получение и обработку данных Госуслуг, ФНС России (данные о налогах и транзакциях), Социального фонда России, МВД России (уголовные и административные дела), ФССП России (данные об исполнительных производствах), Росреестра (включая сведения о движимом и недвижимо-

⁴⁶ Вклад в ВВП России от использования технологий ИИ может достигнуть 2% к 2025 году. Интерфакс (Январь 2023).

⁴⁷ Применение технологий искусственного интеллекта на финансовом рынке (расширенная версия). Ассоциация ФинтТех (Октябрь 2023).

мом имуществе в собственности), ЕГРЮЛ, ЕГРИП, системы ЖКХ (платежи), ЗАГС (информация о семейном положении), ГИБДД России, судов (иски на рассмотрении) и других. Указанные данные позволят финансовым организациям повысить эффективность кредитного скоринга потенциальных заемщиков, предлагать им более персонализированные и качественные продукты и услуги. В случае создания таких наборов данных и предоставления к ним доступа у участников рынка должны быть равные возможности по их использованию. Это придаст импульс развитию ИИ, обеспечит необходимые конкурентные условия.

Вместе с тем следует учитывать, что такие сведения могут содержать персональные данные, банковскую тайну, налоговую тайну и иные виды информации, охраняемой законом. Таким образом, порядок их обработки, а также возможность и условия предоставления участникам финансового рынка требуют детальной проработки со стороны регуляторов в целях недопущения нарушения прав и законных интересов клиентов.

Одним из аспектов развития решений в области ИИ является создание условий доступа участников рынка к государственным данным с учетом обеспечения надлежащего уровня безопасности и защиты прав граждан. Так, Национальная стратегия развития искусственного интеллекта на период до 2030 года⁴⁸ в качестве одного из направлений указывает на необходимость повышения доступности и качества данных, востребованных для развития ИИ, и в том числе государственных. Концепция развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники на период до 2024 года⁴⁹ предлагает совершенствовать режим оборота данных посредством в том числе создания правовых условий для доступа разработчиков систем ИИ к данным, собираемым государственными органами. В качестве необходимого шага указывается расширение практики раскрытия данных, содержащихся в информационных ресурсах органов и организаций государственного сектора, создание основ правового регулирования функционирования общедоступных платформ таких данных.

При участии Банка России продолжается реализация мероприятий по развитию инфраструктуры Цифрового профиля – единого окна по обмену данными, необходимыми для предоставления финансовых продуктов и услуг, между гражданами, государством и бизнесом. Совместно с уполномоченными федеральными органами исполнительной власти Банк России разрабатывает регулирование, направленное на расширение перечня сведений, их актуализацию и расширение видов и количества финансовых организаций, использующих Цифровой профиль гражданина. Созданы правовые основы для развития Цифрового профиля юридического лица и предоставления большего количества данных о юридических лицах с целью оказания им дистанционных финансовых услуг.

Развитие механизмов обработки персональных данных и других сведений, относящихся к различным видам тайн

К данным, обрабатываемым финансовыми организациями, применяются различные режимы обработки и ограничения оборота: банковский, страховой, а также общие правила обработки персональных данных. Одной из потребностей организаций финансового рынка для развития технологий ИИ является создание правовых условий оборота обезличенных персональных данных. Под обезличиванием понимается один из способов обработки персональных данных, в результате которого становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту.

В настоящее время вопрос оборота обезличенных данных содержит ряд правовых пробелов. Так, положения Федерального закона от 27.09.2006 № 152-ФЗ «О персональных данных» не предусматривают порядок обработки обезличенных данных. Требования и методы по обезличиванию

⁴⁸ Утверждена Указом Президента Российской Федерации от 10.10.2019 № 490.

⁴⁹ Утверждена Распоряжением Правительства Российской Федерации от 19.08.2020 № 2129-р.

персональных данных утверждены приказом Роскомнадзора от 05.09.2013 № 996⁵⁰. Сфера применения данного приказа формально ограничена государственными и муниципальными органами. Отсутствие регулирования ограничивает возможности развития использования коммерческими организациями обезличенных данных.

В центре отдельного внимания Банка России также находится вопрос аутсорсинга, связанный с применением информационных технологий и облачных сервисов. Основными проблемными зонами выступают защита конфиденциальной информации, содержащей сведения, составляющие охраняемую законом тайну, а также обеспечение операционной надежности технологических процессов, переданных на аутсорсинг или реализуемых с применением облачных сервисов.

Технологии открытых программных интерфейсов для передачи данных (Открытые API)

Использование Открытых API создает новые возможности для оборота данных, одновременно позволяя повысить доступность финансовых услуг для клиентов и обеспечивая возможность управления собственными данными.

Передача данных через Открытые API упрощает процесс взаимодействия между поставщиком и потребителем данных на финансовом рынке и сокращает временные издержки при поиске источников данных. Кроме того, создание платформы согласий обеспечит возможность удобной обработки согласий клиентов на использование данных, что, в свою очередь, повысит доступность данных, в том числе для целей развития ИИ.

Развитие и повышение доступности вычислительной инфраструктуры и программного обеспечения для решения задач в области искусственного интеллекта

Формирование высокопроизводительной вычислительной инфраструктуры для разработки решений в области ИИ

Высокая ресурсоемкость разработки и обучение современных моделей ИИ, особенно генеративного ИИ, обуславливает необходимость значительных инвестиций в вычислительную технику. Кроме того, особенность устройства памяти компьютера ограничивает виды решений, на основе которых возможно эффективно проводить разработку и обучение. Например, для глубокого обучения повсеместно используются графические процессоры⁵¹.

В связи с этим возникает потребность разработки специализированных кластеров и центров хранения данных непосредственно под задачи ИИ. Один из аспектов развития заключается в необходимости формирования и развития кластеров ИИ. Кластеры выступают как центры притяжения для инноваций, капиталовложений и специалистов. Для реализации этого потенциала требуется скоординированное взаимодействие всех заинтересованных участников, включая государственные органы, бизнес и научное сообщество⁵².

Использование программного обеспечения, распространяемого на основе открытых лицензий

Под Open Source подразумевается программное обеспечение с открытым исходным кодом, представляющее собой компьютерную программу, исходный код которой является общедо-

⁵⁰ Приказ Роскомнадзора от 05.09.2013 № 966 «Об утверждении требований и методов по обезличиванию персональных данных».

⁵¹ [Frontier AI Regulation: Managing Emerging Risks to Public Safety](#). Anderljung, Barnhart, Korinek, Leung, O'Keefe, Whittlestone (September 2023).

⁵² [A blueprint for building national compute capacity for artificial intelligence](#). OECD Digital Economy Papers (February 2023).

ступным для использования, модификации и распространения на основании специального лицензионного соглашения, содержащего минимальное количество ограничений. Благодаря этому многие индивидуальные разработчики, стартапы и учебные заведения могут исследовать, разрабатывать и применять ИИ без необходимости лицензировать дорогостоящее коммерческое программное обеспечение. Так как исходный код доступен всем, разработчики могут быстро адаптировать и улучшать существующие алгоритмы, что способствует быстрому развитию технологий ИИ.

Вместе с тем использование программного обеспечения с открытым исходным кодом (Open Source) в настоящее время сопровождается рядом рисков, связанных с возникновением инцидентов информационной безопасности. Например, такой программный код может содержать вредоносный код и другие недекларируемые возможности (например, лозунги). Несмотря на доступность и удобность решений Open Source, их применение должно сопровождаться повышенным вниманием со стороны пользователей и разработчиков.

Развитие отечественного программного обеспечения, используемого для обучения ИИ

Реализация потребностей участников российского финансового рынка в технологиях сталкивается с препятствиями из-за ограничений, связанных с введением санкций, ограничивающих доступ к зарубежным решениям. По данным опроса Ассоциации ФинТех, 40% российских компаний обеспокоены незначительным количеством отечественных инструментов для работы с машинным обучением. По данным ряда СМИ, на фоне реализации геополитических рисков стоимость разработки программного обеспечения и ИТ-систем на базе ИИ за год увеличилась в России на 30–40% в 2022 году.

Развитие отечественных решений в сфере информационно-коммуникационных технологий⁵³ – одна из целей развития российской экономики в среднесрочной перспективе. Отечественные ИТ-решения усилят технологический суверенитет, позволив финансовым организациям быстрее адаптироваться к особенностям российского рынка и обеспечить устойчивость отечественной вычислительной (технологической) инфраструктуры.

Развитие научной базы в области искусственного интеллекта

В настоящее время скорость интеграции новых решений при разработке и развитии традиционного и генеративного ИИ определяет дальнейшую жизнеспособность моделей. Результаты исследований, публикуемые в научных работах, в кратчайшие сроки применяются разработчиками ИИ в своих продуктах.

Для поддержания темпов развития ИИ необходимо существенное повышение объема исследований и разработок в области ИИ.

Россия обладает существенным потенциалом для того, чтобы стать одним из международных лидеров в развитии технологий ИИ. Национальной стратегией развития искусственного интеллекта на период до 2030 года⁵⁴ поддержка научных исследований в целях обеспечения опережающего развития ИИ указана среди приоритетных направлений деятельности.

Предполагается, что в рамках реализации указанной стратегии к 2024 году Россия войдет в топ-10 стран по количеству статей и участию в конференциях, а к 2030 году – в топ-10 стран по среднему уровню цитируемости.

Систематическая поддержка российских научных исследований позволит сформировать необходимую основу для внедрения передовых решений практических задач.

⁵³ Согласно подп. «д» п. 1, подп. «д» п. 2 Указа Президента Российской Федерации от 21.07.2020 № 474 «О национальных целях развития Российской Федерации на период до 2030 года».

⁵⁴ Утверждена Указом Президента Российской Федерации от 10.10.2019 № 490.

Развитие профессиональных компетенций в области искусственного интеллекта

Одно из ключевых условий развития ИИ на финансовом рынке – **обеспечение отрасли квалифицированными кадрами**. Согласно результатам опроса Ассоциации ФинТех, недостаток профильных специалистов по ИИ является ключевым барьером внедрения ИИ, наличие данной проблемы отметили 84% участников исследования. Респонденты выделяют три категории наиболее востребованных специалистов: дата-сайентисты, дата-инженеры и дата-аналитики, при этом наибольший дефицит наблюдается в привлечении квалифицированных дата-инженеров⁵⁵.

Для обеспечения потребностей рынка необходимы взаимодействие школ, университетов и рынка, подготовка и переподготовка специалистов на базе университетов и крупных технологических компаний. Примером организации подготовки студентов и переподготовки опытных специалистов является работа Финтех Хаба Банка России, направленная на создание дополнительных к основному образованию учебных программ, нацеленных на приобретение практических навыков, необходимых для создания проектов на основе ИИ для финансовой индустрии.

Повышение доверия граждан к технологиям искусственного интеллекта

Одним из ключевых условий для развития ИИ на финансовом рынке является **доверие граждан к технологии ИИ**. Согласно опросу ВЦИОМ, 87% россиян в той или иной степени знают о технологии ИИ, в то время как доверяют технологии ИИ 55% респондентов. Основные мотивы доверия: возможность передать ИИ опасные для человека виды работ, объективность ИИ (безэмоциональность, отсутствие личных интересов), улучшение и упрощение жизни и работы человека. Главная причина недоверия связана с тем, что, по мнению респондентов, в работе ИИ могут возникать ошибки и сбои⁵⁶.

Проект Стратегии повышения финансовой грамотности и формирования финансовой культуры до 2030 года, подготовленный Минфином России и Банком России, направлен в том числе на повышение у населения уровня финансовой киберграмотности – основных знаний, умений и навыков граждан в области безопасного пользования цифровыми и платежными технологиями. Результатом должно стать среди прочего умение защищать права потребителя с помощью правовых инструментов в цифровой среде, понимание и соблюдение правил финансовой кибербезопасности при использовании цифровых и платежных технологий, умение осмотрительно относиться к размещению в Интернете личных данных. Проект стратегии имеет широкую сферу действия и может распространяться также и на финансовые продукты, в которых используется технология ИИ.

Кроме того, федеральный проект «Искусственный интеллект» предусматривает реализацию ряда мероприятий, направленных на повышение доверия к ИИ.

Вместе с тем основополагающее влияние на повышение доверия граждан к применению технологий ИИ на финансовом рынке, как и в случае внедрения иных инноваций, будет оказывать накопление гражданами качественного пользовательского опыта, связанного с оказанием финансовыми организациями клиентам услуг с использованием ИИ без возникновения у клиентов затруднений, а также безопасность применения технологий ИИ.

⁵⁵ [Применение технологий искусственного интеллекта на финансовом рынке](#). Ассоциация ФинТех (Октябрь 2023).

⁵⁶ [Искусственный интеллект: угроза или светлое будущее?](#) ВЦИОМ (Декабрь 2022).

Развитие конкуренции

Даже в условиях равного доступа к данным все равно могут возникать проблемы, связанные с опережающим развитием ИИ у крупных технологических игроков, которые обладают более серьезными возможностями для инвестиций в инновации.

По оценке Ассоциации ФинТех, лидеры рынка финтеха в России за последние 10 лет инвестировали в решения на основе ИИ около 600 млрд рублей. При этом средние и небольшие российские финансовые компании инвестируют в ИИ в среднем в 500 раз меньше, чем крупные организации: от 100 до 300 млн рублей в год⁵⁷.

С учетом объективных факторов (отсутствуют необходимые средства на капитальные вложения, финансовая отдача носит долгосрочный и в ряде случаев неопределенный характер, невозможно привлечь компетентный персонал и так далее), препятствующих активному внедрению инструментов ИИ небольшими компаниями, возможны следующие меры стимулирования более широкого применения ИИ на рынке:

- предоставление целевой финансовой поддержки организациям, осуществляющим деятельность по развитию и внедрению технологий ИИ в рамках соответствующей стратегии⁵⁸;
- льготные условия привлечения финансирования проектами по разработке ИИ в рамках Таксономии приоритетных проектов технологического суверенитета и структурной адаптации экономики⁵⁹;
- создание благоприятных регуляторных условий для использования финансовыми организациями аутсорсинга решений в области ИИ, включая решения, разработанные независимыми технологическими компаниями или новыми участниками рынка ИИ (в том числе основанными при совместном участии участников рынка, государства и инвесторов).

⁵⁷ [Применение технологий искусственного интеллекта на финансовом рынке \(расширенная версия\)](#). Ассоциация ФинТех (Октябрь 2023).

⁵⁸ Создание стимулов для привлечения инвестиций, а также поддержка экспорта, в частности, предусмотрены п. 25 «Национальной стратегии развития искусственного интеллекта на период до 2030 года», которая утверждена Указом Президента Российской Федерации от 10.10.2019 № 490.

⁵⁹ Таксономия технологического суверенитета утверждена постановлением Правительства Российской Федерации от 15.04.2023 № 603. Для поддержки финансирования проектов в рамках таксономии предусмотрены льготы для банковского кредитования и определены специальные условия для вложений ПИФ и профессиональных участников рынка ценных бумаг в облигации проектов, соответствующих данной Таксономии.

5. РИСКИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА ФИНАНСОВОМ РЫНКЕ

Внедрение ИИ может оказывать влияние как на бизнес-процессы участников рынка, так и на финансовую систему в целом. Присущие ИИ риск-факторы⁶⁰ (например, проблемы объяснимости алгоритмов и качества применяемых данных) могут значительно усиливать классические риски финансового сектора, в частности приводить к фальсификации клиентских данных, мошенничеству и случаям недобросовестной конкуренции⁶¹.

Обсуждение рисков ведется на разных площадках, в том числе и на зарубежных. Так, в рамках Рабочей группы по ИИ ОЭСР была создана специальная Экспертная группа по рискам и подотчетности в области ИИ⁶², куда вошли представители Международной организации по стандартизации (ISO), Европейской комиссии, Совета Европы, ЮНЕСКО, Института ответственного ИИ (RAII). Аналогичная работа проводится также на площадках Международного валютного фонда⁶³, Банка международных расчетов⁶⁴, Совета по финансовой стабильности⁶⁵. В 2023 году возможности и риски ИИ для международного сообщества также обсуждались Советом Безопасности ООН⁶⁶. В частности, в фокусе внимания регуляторов находятся вопросы применения генеративного ИИ и машинного обучения на основе синтетических данных в финансовом секторе.

Несмотря на то что стандартизированной карты рисков ИИ и единого взгляда на специфику их влияния на финансовый рынок пока не выработано, **в целом усилия экспертов направлены на оценку и изучение следующих риск-факторов (параметров) при использовании моделей ИИ.**

Риски разработки ИИ и искажения работы модели ИИ

Разработчики ИИ создают модели для решения определенных практических задач на основе имеющихся технических, финансовых и организационных возможностей. Применение моделей ИИ может приводить к возникновению рисков, основанных на неверных или неправильно интерпретированных результатах моделей. Реализация данного риска может привести к финансовым потерям, ошибочным решениям, репутационным последствиям.

Источником рисков некорректной разработки моделей и искажения результатов ее работы могут стать следующие факторы.

Модель может содержать фундаментальные ошибки (например, ошибки программного кода), что может привести к некорректным расчетам и неточным прогнозам.

Модель может быть использована не по назначению. В связи с тем что модели ИИ обучаются для решения конкретных задач, применение их для решения иных задач может приводить к ошибочным результатам работы.

Данные, которые использует модель в процессе эксплуатации, могут существенно отличаться по статистическим свойствам от данных, на которых она была разработана. Неточные и неполные данные могут исказить процесс выявления закономерностей и привести к ошибочным результатам. Например, алгоритмическая система распознавания лиц, обученная исключительно на основе лиц людей определенной расы, пола или возраста, может испыты-

⁶⁰ Факторы, которые увеличивают вероятность реализации риска.

⁶¹ Edmund L. Andrews. [How Unique Are the Risks Posed by Artificial Intelligence?](#) (December 2021).

⁶² [Expert Group on AI Risk & Accountability](#). OECD.AI community.

⁶³ [Generative Artificial Intelligence in Finance: Risk Considerations](#). IMF (August 2023).

⁶⁴ [Humans keeping AI in check – emerging regulatory expectations in the financial sector](#). BIS (August 2021).

⁶⁵ [Artificial intelligence and machine learning in financial services](#). FSB (November 2017).

⁶⁶ [Artificial Intelligence: Opportunities and Risks for International Peace and Security – Security Council, 9381st Meeting](#). UN (July 2023).

вать трудности при обработке лиц людей других рас, полов и возрастных групп. Искажения также могут возникать и вследствие использования некорректных данных. Поскольку большие данные, используемые в ИИ, среди прочего включают сведения из открытых источников, могут множиться ошибки, содержащиеся в этих данных.

Недостаток информации в обучающем наборе данных, а также их качество могут нарушить точность предсказаний и вызвать нестабильность работы модели.

Дополнительные источники риска, связанные с качеством работы моделей, возникают на фоне развития генеративных моделей ИИ. Например, известен феномен так называемых «галлюцинаций» ИИ. Причины их возникновения до конца не изучены, однако отмечается, что такие факторы, как значительные пробелы или многозначные слова в массиве данных для обучения модели, могут приводить к подобным ошибкам⁶⁷.

Зарубежные исследователи в области машинного обучения также отмечают, что генеративные модели могут быть подвержены риску «смещения» – ситуации, когда функциональные способности модели при решении отдельных задач регрессируют с течением времени⁶⁸. Использование синтетических данных для обучения и тестирования моделей также связано с рядом трудностей. Так, их качество зависит от алгоритма, создающего такие данные, а возможность воспроизведения присущих реальности допущений и пробелов может привести к ошибочным результатам обучения и тестирования ИИ.

Риски аутсорсинга (связаны с вмешательством в бизнес-процессы и делегированием отдельных функций при привлечении третьих лиц)

Организации могут прибегать к решениям ИИ, предлагаемым сторонними поставщиками (в том числе с использованием внешних баз данных и облачных вычислений). Это может привести к соответствующим рискам: концентрации поставщиков, нарушению конфиденциальности обрабатываемых данных, сложностям в интерпретации результатов работы ИИ, операционным рискам и так далее. Банк России подробно освещал вопросы риска аутсорсинга на финансовом рынке, в том числе в сфере информационных технологий и облачных сервисов, в одноименном докладе для общественных консультаций⁶⁹.

Риски в области данных и информационной безопасности

Согласно результатам опроса Ассоциации ФинТех, 52% компаний, внедряющих и разрабатывающих ИИ-решения, отмечают риски в области данных и **информационной безопасности** в качестве наиболее значимых рисков внедрения ИИ⁷⁰. Одним из рисков внедрения ИИ в сфере информационной безопасности, который вызывает беспокойство, является **риск утечки персональных данных** пользователей, а также сведений, составляющих отдельные виды тайн (банковскую, налоговую, страховую и другие), которые могут содержать наборы данных, сформированные для обучения моделей ИИ. В случае кражи данных, на которых обучался ИИ, возможно последующее частичное или полное воспроизведение алгоритма ИИ, в результате чего владелец ИИ может понести репутационные издержки.

Кроме того, системы ИИ подвержены **рискам мошеннических атак** на стадии как обучения (атаки данных), так и эксплуатации моделей (атаки моделей). Изменение данных приводит к ухудшению результативности работы алгоритмов машинного обучения и нарушению работы модели.

⁶⁷ [Можно ли раз и навсегда устранить «галлюцинации» искусственного интеллекта?](#) Sber AI (Сентябрь 2023).

⁶⁸ L. Chen, M. Zaharia, J. Zou [How is ChatGPT's Behavior Changing over Time?](#) (July 2023).

⁶⁹ [Управление рисками аутсорсинга на финансовом рынке](#). Банк России (2022).

⁷⁰ [Применение технологий искусственного интеллекта на финансовом рынке](#). Ассоциация ФинТех (Октябрь 2023).

Ключевая проблема атак данных в том, что для модели измененные данные являются такими же легитимными, как и реальные данные, необходимые для обучения и проверки модели⁷¹. В связи с этим нет явного способа установить достоверную причину ухудшения качества данных.

Этические риски и риски нарушения прав потребителей и инвесторов⁷²

В международной практике к этическим рискам причисляют проблемы, связанные с предвзятостью, дискриминацией, выявлением ложных корреляций и непрозрачными и непонятными решениями ИИ.

Дискриминация и индивидуальное ценообразование. С помощью ИИ организации могут обрабатывать растущий объем данных о потребителях и их характеристиках (потребительские предпочтения и поведение, лояльность к бренду и прочее). Это позволяет участникам не только устанавливать цены для широкой аудитории потребителей, но и в некоторых случаях индивидуально адаптировать цены на основе готовности потребителя платить⁷³. Также такие системы могут учитывать при ценообразовании финансовых продуктов и оценке рисков отдельных клиентов дискриминационные факторы (вероисповедание, этническая принадлежность и так далее).

Пример предвзятой кластеризации искусственного интеллекта

Важность недискриминационного применения ИИ при оценке кандидатов подчеркивается примером компании Amazon. В августе 2018 года компании пришлось отказаться от программы найма сотрудников с использованием ИИ, так как выяснилось, что она исключала женщин из списка кандидатов на рассмотрение и иногда нанимала менее квалифицированных работников-мужчин.

Основываясь на исторических данных о приеме на работу за последние 10 лет в технологической отрасли, где раньше доминировали мужчины, алгоритм выделял в резюме соискателей слова, указывающие на членство в женских клубах, отсеивал или понижал рейтинг выпускниц колледжей и школ для женщин.

Ограниченная способность решать сложные запросы и применение не подходящих для обслуживания потребителей программ. Чат-боты полезны для выполнения простых задач (запросить кредитный рейтинг, оспорить транзакцию, получить информацию о продукте), но их эффективность снижается по мере усложнения клиентских запросов. Так, если алгоритм не может идентифицировать запрос или цель действий клиента, то, как правило, направляет ему базовую информацию или отсылку к разделу часто задаваемых вопросов (FAQ), повторяя цикл снова и снова. Некоторые алгоритмы не способны различать личностные особенности речи или иностранный язык, что также может привести к дискриминации уязвимых слоев населения.

Нарушение интересов потребителей и инвесторов при предложении финансовых продуктов и услуг. ИИ служит инструментом для оптимизации прибыли финансовых организаций, анализируя привычки потребления контента, товаров и услуг у определенных целевых групп. Интересы компаний владельцев ИИ могут быть как явно, так и неявно встроены в оптимизационные механизмы ИИ-систем, обучающихся на основе примеров поведения пользователей. При этом могут не учитываться интересы отдельных потребителей, в том числе тех, которые могут не обладать знаниями в области ИИ и (или) не располагать информацией о применении таких технологий. Например, чат-боты, используемые для рассмотрения запросов и жалоб клиентов, которые могут не осознавать, что имеют дело с автоматизированной системой, могут вводить определенные сегменты клиентов в заблуждение в силу предубеждений, сформировавшихся на стадии обучения моделей, лежащих в основе таких чат-ботов.

⁷¹ Намиот Д.Е. [Схемы атак на модели машинного обучения](#) (2023).

⁷² [The ethics of artificial intelligence: Issues and initiatives](#). European Parliament (2020).

⁷³ Т. Ван. Человеческое понимание, отсутствующее в больших данных (2016).

Некорректные способы продажи и продвижения продуктов, не отличимые от взаимодействия с сотрудниками финансовой организации, также могут спровоцировать потребителя сделать выбор в пользу менее подходящих ему услуг и увеличить его транзакционные издержки.

Роботизированные сервисы, используемые компаниями для обслуживания массового инвестора на фондовом рынке, могут скрывать в себе риски продвижения собственного инвестиционного продукта; побуждать совершать излишнюю ребалансировку портфеля, приводя к дополнительным расходам для пользователя; приводить к пассивности поведения инвестора на рынке в части оценки принимаемых им инвестиционных решений, в том числе за счет элементов геймификации, автоматизированности процесса инвестирования.

Таким образом, особое внимание должно уделяться оценке применения технологий ИИ (как моделей и алгоритмов в более широком плане) с точки зрения обеспечения должной защиты прав потребителей финансовых услуг и инвесторов.

Риски конкуренции

Количество участников, вид услуг и другие общие характеристики рынка, на котором используется ИИ, во многом определяют те виды рисков, которым может быть подвержена конкуренция: начиная от содействия явному или молчаливому сговору⁷⁴ и заканчивая рисками слияний и поглощений, связанных с технологиями ИИ. Применение ИИ в отдельных случаях может усиливать и помогать по-новому использовать уже известные антиконкурентные практики.

ИИ и сговор

- 1) Технологии ИИ используют данные о потребителях, предложения конкурентов и операции на рынке. Доступность данных о ценах конкурентов или операциях при высокой степени прозрачности рынка делает сговор более вероятным. Организации, которые будут стремиться к созданию картеля или заключению картельного соглашения, для связи друг с другом могут использовать косвенный способ – сигнализирование. Алгоритмы могут определять сигналы, действовать в соответствии с ними по заранее установленным правилам принятия решений и отслеживать факт принятия приглашения к сговору потенциальными членами картеля.
- 2) На рынках частого взаимодействия с конкурентами использование инструментария ИИ упрощает выполнение и отслеживание участниками соглашений о сговоре (например, предоставляется возможность более быстрой корректировки цен). Более того, алгоритм помогает избежать неверного толкования или ошибок в реализации картельных соглашений на основе общего потока данных, применяя заранее установленные параметры. Более сложные приложения на основе ИИ используются для мониторинга реализации, выявления отклонений от соглашений о сговоре, вплоть до применения наказания в случае подтверждения нарушения соглашения.
- 3) Внедряемые технологии ИИ в результате обучения могут принять решение, что использование молчаливого сговора будет эффективной стратегией, в то время как изначально такая модель поведения не закладывалась.

⁷⁴ Явный сговор – антиконкурентное поведение, которое подтверждается явными соглашениями (письменными или устными). Молчаливый сговор – форма антиконкурентной координации, которая может быть достигнута без какой-либо необходимости в явном соглашении, но которое участники рынка могут поддерживать, признавая свою взаимозависимость.

Применение корректирующих алгоритмов

В ноябре 2021 года Суд общей юрисдикции Европейского союза оставил в силе решение Европейской комиссии о штрафе в размере 2,42 млрд евро компании Google за нарушение антимонопольных правил ЕС.

Решением от 27 июня 2017 года Европейская комиссия установила, что компания Google злоупотребила своим доминирующим положением на рынке поисковых онлайн-сервисов в 13 странах Европейской экономической зоны. Алгоритм Google отдавал предпочтение собственным сервисам сравнительных покупок, выделяя их в специальных полях, в то время как к результатам конкурентов применялись корректирующие алгоритмы, понижающие их рейтинг в общем поиске.

Трафик, генерируемый общей поисковой системой Google, также повлиял на итоговые показатели сервисов компании. Обучившись на некорректных данных, алгоритм необоснованно повышал релевантность продуктов Google и снижал трафик конкурирующих с ними сервисов¹.

¹ Google Shopping: General Court of the European Union (Press Release № 197/21, [Case T-612/17](#)). European Commission. [Competition Case AT.39740](#).

Злоупотребление доминирующим положением

На сегодняшний день финансовые организации имеют разные возможности для внедрения в свои бизнес-процессы решений на основе ИИ. Их разработка – капиталоемкий процесс, требующий значительных инвестиций в инфраструктуру и наличия интеллектуальных ресурсов на длительном горизонте времени.

Крупные участники, которые располагают достаточными возможностями для внедрения ИИ, могут оказаться в более выигрышных условиях. Использование ими технологии ИИ, наряду с сильными сетевыми эффектами с учетом проблем дифференциации доступа к данным и переключения потребителя, может стать источником конкурентного преимущества и барьером для входа, способствовать развитию рыночного господства и потенциального доминирования, а также привести к более существенным ограничениям для распространения технологий.

Макроэкономические риски и риски финансовой стабильности⁷⁵

В настоящее время масштаб и характер применения технологий ИИ на финансовом рынке не представляют собой источники системного риска. Тем не менее при росте использования технологии ИИ они могут привести к эффекту накопления рисков и нестабильности на финансовом рынке:

1. Системам ИИ могут быть свойственны риски, связанные с непрозрачностью результатов (решений), подверженностью манипуляциям, проблемами надежности и конфиденциальности. Их реализация может негативно отражаться на доверии общества к целостности и безопасности финансовой системы.
2. Использование стандартизированных алгоритмов и моделей от узкого круга разработчиков может привести к большей однородности и единообразию – например, в оценке рисков и принятии решений о кредитовании. Вследствие концентрации данных и растущего применения альтернативных данных в ИИ могут возникнуть риск единообразия и «риск вне выборки»⁷⁶, что в конечном счете может привести к негативному системному влиянию на рынок⁷⁷.
3. Поставщики услуг ИИ могут стать системообразующими участниками инфраструктуры финансового рынка из-за высокой специализации систем ИИ и сетевых эффектов, что может увеличить уязвимость финансовой системы к единым «точкам отказа».

⁷⁵ El Bachir Boukherouaa, Khaled AlAjmi, Jose Deodoro, Aquiles Farias and Rangachary Ravikumar. [Opportunities and Risks of Artificial Intelligence in Finance](#). IMF (2021).

⁷⁶ Риск возникновения ошибки из-за использования неподходящих процедур оценки и, как следствие, получения на их основе неверного результата (non-sampling risk).

⁷⁷ El Bachir Boukherouaa, Khaled AlAjmi, Jose Deodoro, Aquiles Farias and Rangachary Ravikumar. [Opportunities and Risks of Artificial Intelligence in Finance](#). IMF (2021).

Модели ИИ могут хорошо работать в относительно стабильной среде, которая генерирует надежные сигналы, позволяя отслеживать тенденции без значительной потери точности прогноза. Однако в кризисных ситуациях, предполагающих значительный поведенческий сдвиг, возможно снижение точности моделей ИИ, что является дополнительным негативным фактором не только для участников рынка, применяющих соответствующие модели, но и в целом для финансовой стабильности.

Возможная восприимчивость алгоритмов ИИ к атакам с подложными и некорректными данными может приводить к паническому поведению значительного количества субъектов экономики и нарушать нормальное функционирование финансового рынка.

Широкое использование ИИ потенциально может повысить процикличность на финансовом рынке: алгоритм автоматизирует процесс, тем самым его ускоряя. Аналогичное негативное влияние на финансовый рынок может оказать однотипность используемых алгоритмов, поскольку в совокупности это может привести к проциклическому поведению многих участников.

Дополнительно в части рисков использования генеративного ИИ, по мнению экспертов, выделяются следующие риски.

1. Объяснимость результатов

Модели генеративного ИИ работают на основе глубоких нейросетей, обучаемых на больших объемах разных данных, используют множество параметров для генерации ответов и, как следствие, обладают худшей объяснимостью.

2. Стадное поведение (*herd mentality*)

Решения финансовых организаций на основе сгенерированных ИИ отчетов о состоянии экономики, обзоров рынков или рисков в кризисных ситуациях могут оказаться ошибочными и масштабироваться (эффект стадного поведения), особенно если модели генеративного ИИ обучались на данных, собранных в благоприятные для рынка периоды.

3. Риски ликвидности

Алгоритмы генеративного ИИ могут применяться при совершении сделок с высоким кредитным или рыночным риском с целью максимизации прибыли при неправильном обучении моделей в области управления рисками.

Ложный и вредоносный контент, созданный генеративным ИИ, может вызвать панику среди населения. В случае финансовых услуг это может привести, например, к бегству вкладчиков (*bank run*)⁷⁸.

⁷⁸ [Generative Artificial Intelligence in Finance: Risk Considerations](#). IMF (August 2023).

6. ПОДХОДЫ К РЕГУЛИРОВАНИЮ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

6.1. Зарубежный опыт регулирования искусственного интеллекта

Эксперты отмечают, что [в ряде юрисдикций](#) ИИ регулируется при помощи сочетания инструментов:

- жесткого регулирования (hard-law) – юридически обязательных норм, разрабатываемых законодателями и регуляторами;
- мягкого регулирования (soft-law) – не обязательных, но предназначенных для оказания нормативного воздействия, издаваемых и поощряемых регулятором норм (рекомендательные письма, стандарты);
- саморегулирования (self-regulation) – не обязательных к применению стандартов, принципов и кодексов поведения, созданных участниками рынка без непосредственного участия регулятора⁷⁹.

При этом ряд экспертов отмечают, что страны могут делать акценты на различных инструментах регулирования в отношении ИИ, а подходы к регулированию ИИ варьируются в различных юрисдикциях⁸⁰.

На данный момент условно можно выделить **три основных подхода к регулированию ИИ:**

1. **Ограничительный подход** предусматривает принятие норм, регламентирующих разработку, внедрение и эксплуатацию систем ИИ, содержащих прямой запрет на использование отдельных систем ИИ, а также дополнительные требования к эксплуатации ИИ с повышенным риском. По мнению экспертов, данный подход реализуют *Европейский союз* и *Бразилия*.
2. **Гибридный подход** представляет собой сочетание инструментов жесткого регулирования, мягкого регулирования и саморегулирования на основе риск-ориентированных принципов. К такому подходу можно отнести [модели регулирования](#) таких стран, как *Китай*, *Канада* и *США*.
3. **Стимулирующий подход** предполагает применение инструментов мягкого регулирования (саморегулирование, этические принципы) и полное отсутствие ограничительных мер в отношении ИИ. Примеры данного подхода к регулированию – *Великобритания* и *Сингапур*.

Стоит также отметить, что на фоне широкого распространения генеративного ИИ регулирование таких моделей становится актуальным вопросом для многих юрисдикций. Отдельные инициативы по регулированию ИИ появились в Китае, нормы, регулирующие внедрение и использование генеративных моделей ИИ, содержатся в проектируемом регулировании ИИ в Европейском союзе.

В связи с тем что развитие технологий ИИ сопряжено с рядом рисков технологического, экономического и этического характера, а также с рисками в области информационной безопасности, основной вектор развития регулирования ИИ – применение **принципов риск-ориентированного регулирования**, которое предполагает зависимость уровня регуляторных требований от объема и вероятности реализации рисков внедрения технологий ИИ.

⁷⁹ Wojtek Buczynski, Felix Steffek. [Hard Law and Soft Law Regulations of artificial Intelligence in Investment Management](#) (2022).

⁸⁰ Подробнее см. Wojtek Buczynski, Felix Steffek. [Hard Law and Soft Law Regulations of artificial Intelligence in Investment Management](#) (2022); [Japan's Approach to AI Regulation and Its Impact on the 2023 G7 Presidency](#). Center for Strategic & International Studies (February 2023); Gary Marchant, Carlos Ignacio Gutierrez. [Soft Law 2.0: An Agile and E Soft Law 2.0: An Agile and Effective Governance Appr ernance Approach for oach for Artificial Intelligence](#) (June 2023); Rolf H. Weber. [Artificial Intelligence ante portas: Reactions of Law](#) (August 2021); Alex Engler. [The EU and U.S. diverge on AI regulation: A transatlantic comparison and steps to alignment](#) (April 2023).

Инициативы международных организаций в сфере регулирования искусственного интеллекта

Один из основополагающих документов – опубликованные в 2019 году рекомендации ОЭСР, которые содержат пять принципов использования ИИ:

1. Содействие инклюзивному росту и устойчивому развитию.
2. Ориентированность на человека, добросовестность.
3. Прозрачность и интерпретируемость моделей ИИ.
4. Надежность и безопасность систем ИИ.
5. Подотчетность провайдеров ИИ⁸¹.

Другой пример верхнеуровневых принципов – [Рекомендации об этических аспектах ИИ](#) Генеральной конференции ЮНЕСКО по вопросам образования, науки и культуры 2021 года. Критерии, по которым необходимо оценивать технологии ИИ (в том числе методологии обучения интеллектуальных систем, автоматизированного рассуждения и иные параметры), должны отвечать концепциям права, открытости, доступности и многостороннего участия⁸².

Рекомендации ОЭСР и ЮНЕСКО являются примерами мягкого регулирования, на основе которого многие страны разрабатывают собственные подходы и правила. Так, основываясь на рекомендациях ОЭСР в области применения технологий ИИ, в октябре 2023 года страны G7 согласовали 11 руководящих принципов (Guiding Principles) для разработчиков передовых систем ИИ (Advanced AI Systems), а также **кодекс поведения** (Code of Conduct) разработчиков таких систем⁸³. **В документе подчеркивается, что принципы применимы к моделям генеративного ИИ.** Принципы носят рамочный характер, и отмечается, что юрисдикции могут использовать собственные уникальные подходы к их реализации.

⁸¹ [Recommendation of the Council on Artificial Intelligence](#). OECD (2019).

⁸² [Recommendation on the Ethics of Artificial Intelligence](#). UNESCO (November 2021).

⁸³ [Hiroshima Process International Code of Conduct for Advanced AI Systems](#) (October 2023).

Руководящие принципы G7 для разработчиков передовых систем ИИ¹

Документ призывает разработчиков передовых систем ИИ следовать основным **руководящим принципам**:

1. Принимать адекватные меры по выявлению, оценке и снижению **рисков**, связанных с технологией ИИ, на протяжении всего жизненного цикла систем.
2. Отслеживать уязвимости, возникающие риски и примеры неправомерного использования систем ИИ после их выпуска на рынок.
3. Обнародовать информацию о возможностях, ограничениях, областях применения систем и возможностях их неправомерного использования.
4. Производить обмен информацией об инцидентах, связанных с использованием ИИ, с другими разработчиками ИИ, компаниями, государственными органами, представителями научного сообщества и гражданского общества.
5. Разрабатывать, внедрять и раскрывать политику по управлению рисками, основанную **на риск-ориентированных принципах**.
6. Внедрять надежные средства обеспечения **физической безопасности и кибербезопасности** на всех этапах жизненного цикла систем ИИ.
7. Разрабатывать и внедрять надежные механизмы **аутентификации и подтверждения происхождения контента** (водяные знаки), созданного с помощью **генеративного ИИ**.
8. В приоритетном порядке проводить научные исследования и осуществлять инвестиции в мероприятия, направленные на управление социальными рисками и рисками безопасности систем ИИ.
9. Уделять повышенное внимание разработке систем ИИ, направленных на решение глобальных проблем (климат, здравоохранение, образование).
10. Содействовать разработке и принятию международных технических стандартов.
11. Реализовывать адекватные меры по обеспечению качества данных для ИИ, защите персональных данных и интеллектуальной собственности.

В развитие указанных принципов страны G7 согласовали **Кодекс поведения разработчиков передовых систем ИИ**. Кодекс носит **добровольный характер**.

Данные принципы могут относиться ко всем областям ИИ, включая разработку, развитие, продвижение или практическое применение систем ИИ.

¹ [Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System](#). G7 (October 2023).

Руководство IOSCO для рыночных посредников и управляющих активами при использовании ИИ¹

IOSCO предлагает следующие требования для финансовых организаций:

1. Назначить представителя высшего менеджмента, ответственного за надзор над разработкой, тестированием, применением и контролем за технологиями ИИ. Такие обязанности может выполнять лицо, осуществляющее наблюдение за применением общих технологий и данных в компании. Организациям рекомендуется иметь задокументированную внутреннюю систему управления с четко определенными линиями подотчетности.
2. Проводить тестирование и мониторинг технологий ИИ, чтобы убедиться, что алгоритмы: 1) ведут себя ожидаемым образом как в обычных рыночных условиях (unstressed), так и в условиях напряженности на рынке (stressed market conditions); 2) работают в соответствии с регуляторными требованиями.
3. Убедиться, что ответственные сотрудники обладают необходимыми навыками и опытом для разработки, тестирования, внедрения, мониторинга и надзора за применяемыми ИИ, а также способны проводить комплексную оценку стороннего провайдера, включая уровень его знаний, компетенций и опыта.
4. Понимать степень зависимости от сторонних поставщиков технологий ИИ. Рекомендуется иметь соглашение об уровне обслуживания (service level agreement), определяющее перечень переданных на аутсорсинг функций, ответственность поставщика услуг и, если возможно, показатели эффективности предоставления услуги и санкции за неудовлетворительные результаты работы.
5. Определить тип и уровень раскрытия информации о применяемых организациями технологиях ИИ пользователям услуг и регуляторам.
6. Разработать и утвердить внутренние механизмы контроля, направленные на обеспечение надлежащего качества и объема данных, от которых зависят обоснованность и результат применения ИИ (количество искажений).

В ходе обсуждения одноименного консультационного доклада респонденты IOSCO признали важность формирования нормативно-правовой базы в части применения технологий ИИ и приверженности стандартам добросовестного поведения. В результате опроса также подтвердилась необходимость в четком определении ИИ и применении *пропорционального подхода* в регулировании таких технологий.

¹ [The Use of Artificial Intelligence and Machine Learning by market intermediaries and asset managers](#). IOSCO (2020).

1. Юрисдикции, использующие ограничительный подход к регулированию искусственного интеллекта

Европейский союз

Ключевой элемент регулирования в сфере ИИ – проект Регламента об искусственном интеллекте⁸⁴, который устанавливает гармонизированные правила размещения, ввода в эксплуатацию и использования ИИ в ЕС (AI Act⁸⁵). В соответствии с AI Act предлагается принять общее определение систем ИИ в законодательстве ЕС и применять пропорциональное регулирование на основе *риск-ориентированного подхода*. Модели ИИ предлагается классифицировать в зависимости от уровня создаваемых ими угроз:

- **ИИ с неприемлемым уровнем риска:** запрет на использование ИИ, осуществляющих социальный скоринг и биометрическую идентификацию в режиме реального времени в публичных местах.
- **ИИ с высоким уровнем риска:** требования к уровню надежности, безопасности, качеству обучающих данных и раскрытию информации, а также к наличию процедур оценки и управления рисками. Системы ИИ, используемые для *оценки кредитоспособности физических лиц*, также отнесены законодателями к данной группе, поскольку они могут определять доступ к финансовым ресурсам или основным услугам, например жилью.

⁸⁴ В июне 2023 года регламент был одобрен Европейским парламентом, следующий этап – согласование окончательного варианта документа Европейской комиссией и утверждение в Совете ЕС.

⁸⁵ [Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence \(Artificial Intelligence Act\) and amending certain union legislative acts](#) (June 2021).

- **ИИ с ограниченным уровнем риска:** требования об уведомлении пользователей о взаимодействии с ИИ и возможности от него отказаться.
- **ИИ с минимальным уровнем риска:** инструменты мягкого регулирования, например разработка кодексов этики.

Отдельные требования касаются операторов систем генеративного ИИ, которые должны будут раскрывать информацию о применении ИИ при создании контента и использовании обучающих данных, защищенных авторским правом, а также обучать и проектировать модели ИИ для предотвращения генерирования незаконного контента.

Общественная реакция в ЕС в связи с разработкой регламента

Разработка регламента вызвала общественный резонанс в ЕС, а его отдельные положения (в том числе в части регулирования генеративных систем) стали объектом критики со стороны разработчиков ИИ. По данным СМИ, в июне 2023 года 150 компаний направили в ЕС письмо против законопроекта, в котором заявили, что избыточно жесткое регулирование непропорционально увеличит регуляторные риски для разработчиков, приведет к снижению конкурентоспособности компаний и оттоку инвестиций из европейских проектов¹.

¹ [European companies sound alarm over draft AI law](#). Financial Times (October 2022).

Надзор за соблюдением положений AI Act в части финансовых услуг будет возложен на уполномоченные органы, осуществляющие надзор на финансовом рынке, и, если применимо, на Европейский центральный банк. Дополнительно указывается, что *законодательство ЕС в части правил и требований к внутренней системе управления рисками в финансовых организациях также распространяется на финансовые услуги с применением ИИ.*

Например, финансовые организации, применяющие ИИ для оценки кредитоспособности, должны будут учитывать положения Регламента о цифровой устойчивости (DORA⁸⁶) и внедрить систему внутреннего управления и контроля для управления рисками информационно-коммуникационных технологий. Такая система должна включать стратегию цифровой операционной устойчивости финансовой организации, определять уровень толерантности к риску и подвергаться внутреннему аудиту в соответствии с планом проверок.

В Руководстве Европейской службы банковского надзора (EBA)⁸⁷ по управлению рисками при кредитовании также выработаны общие *правила для управления рисками вне зависимости от применяемой технологии*: банкам надлежит придерживаться надежных и взвешенных стандартов принятия кредитных рисков и процедур управления ими, следовать требованиям по защите прав потребителей и внутреннего контроля в части ПОД/ФТ.

Необходимо отметить, что в действующем законодательстве ЕС есть нормы, регулирующие отдельные случаи использования ИИ. Например, Общий регламент защиты данных (GDPR) гарантирует право пользователя на получение «значимой информации о логике принятия решений» алгоритмических систем, Закон о цифровых услугах (DSA⁸⁸) предусматривает требования прозрачности для рекомендательных систем, используемых цифровыми платформами⁸⁹, а в Основах кибербезопасности при применении ИИ⁹⁰ (FAISP) предусмотрены три уровня управления рисками кибербезопасности: базовый, специфический для систем ИИ, отраслевой при применении ИИ.

⁸⁶ [Regulation \(EU\) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector amending Regulations \(EC\) № 1060/2009, \(EU\) № 648/2012, \(EU\) № 600/2014, \(EU\) № 909/2014 and \(EU\) 2016/1011](#). Положения регламента полностью вступают в силу 17 января 2025 года.

⁸⁷ [Guidelines on loan origination and monitoring](#). EBA (2020).

⁸⁸ [General Data Protection Regulation](#). Regulation EU 2016/679 of the European Parliament and of the Council of 27 April 2016.

⁸⁹ [The EU and U. S. diverge on AI regulation: A transatlantic comparison and steps to alignment](#). Brookings (April 2023).

⁹⁰ [Framework for Artificial Intelligence Cybersecurity Practices](#). Cyber Risk EU Agency for Cybersecurity.

Инициативы национальных финансовых регуляторов ЕС в сфере ИИ

Наряду с развитием общеотраслевого законодательства ЕС в отношении ИИ национальные финансовые регуляторы формируют подходы к регулированию ИИ на финансовом рынке.

Управление пруденциального надзора (ACPR) совместно с *Управлением по финансовым рынкам (AMF)* Франции опубликовало руководство⁹¹, в котором утверждается, что при использовании ИИ при управлении активами и во внутренних моделях по расчету требований к капиталу следует регулярно обновлять данные и проверять их качество. Для этого рекомендуется совмещать сразу несколько контрольных процедур: от традиционных параллельных процессов обработки частей тестовых данных и экспертных оценок до специально разработанных для этого алгоритмических инструментов.

Центральный банк Нидерландов в 2019 году опубликовал доклад (SAFEST⁹²), в котором выделил шесть принципов применения ИИ в финансовом секторе: надежность, ответственность, справедливость, этичность, компетентность, прозрачность.

Федеральный банк Германии (Deutsche Bundesbank) в 2020 году опубликовал доклад об использовании ИИ на финансовом рынке⁹³, в котором указал, что при разработке регулирования ИИ необходимо оценивать действующее, технологически нейтральное законодательство. Предметом надзора являются не технологии ИИ, а риски, которые реализуются при использовании ИИ в каждом конкретном случае.

Федеральное управление по финансовому надзору Германии (BaFin) в 2021 году опубликовало принципы использования алгоритмов ИИ на финансовом рынке⁹⁴. По мнению регулятора, оценка решений на основе ИИ должна носить *пропорциональный, риск-ориентированный и технологически нейтральный характер*. Предварительная проверка решений необходима в *исключительных случаях* – например, если они используются финансовыми организациями для соблюдения установленных требований к капиталу. В числе принципов использования ИИ адекватный риск-менеджмент, отсутствие предвзятости алгоритмов ИИ, законность систем ИИ.

Бразилия

На рассмотрении Сената Бразилии находится законопроект об использовании ИИ, многие положения которого аналогичны проекту Регламента об ИИ Европейского союза. Законопроект предусматривает применение риск-ориентированного подхода, который устанавливает следующую градацию рисков:

- 1) *запрещенные системы* (системы социального скоринга, биометрическая идентификация в публичных местах);
- 2) *системы с высоким уровнем риска* (системы кредитного скоринга, автономные транспортные средства, использование решений ИИ в медицине, образовании и так далее). Структура управления системами с таким уровнем риска должна отвечать дополнительным требованиям: ведение технической документации, проведение тестирований надежности системы, осуществление процедур по увеличению объяснимости алгоритмов⁹⁵.

⁹¹ [Artificial Intelligence; challenges of the financial sector](#). ACPR, AMF (December, 2018).

⁹² [General principles for the use of Artificial Intelligence in the financial sector](#). De Nederlandsche Bank.

⁹³ [The Use of Artificial Intelligence and Machine Learning in the Financial Sector](#). Deutsche Bundesbank (November 2020).

⁹⁴ [Big Data and Artificial Intelligence: Principles for the use of algorithms in decision-making processes](#). BaFin (2021).

⁹⁵ [AI Report and Draft AI Law](#). Senate Committee Brazil (January 2023).

2. Юрисдикции, использующие элементы гибридного подхода к регулированию искусственного интеллекта

Соединенные Штаты Америки

Значимым шагом в направлении регулирования ИИ является опубликованный Администрацией Президента США в конце октября 2023 года указ о безопасном, надежном и заслуживающем доверия ИИ⁹⁶, а также опубликованный в конце 2022 года Билль о правах в эру ИИ (AI Bill of Rights⁹⁷). Последний документ включает следующие принципы: безопасные и эффективные системы, алгоритмическая защита от дискриминации, конфиденциальность данных, уведомление и объяснение, а также человеческие альтернативы, рассмотрение и отступление.

Ранее, в 2020 году, Администрация Президента США рекомендовала федеральным органам власти и профильным регуляторам придерживаться гибкого и стимулирующего подхода к вопросам регулирования ИИ, проводить мероприятия по оценке рисков, выгод и затрат, обеспечивать прозрачность и доверие, безопасность и недискриминацию при внедрении ИИ. В целях реализации данных принципов рекомендовалось использовать в том числе инструменты мягкого регулирования: 1) тематические руководства; 2) проведение экспериментов и создание пилотных программ; 3) добровольные стандарты.

В качестве ключевых регуляторных практик можно назвать следующие документы:

- *Руководство для разработчиков и пользователей ИИ* Национального института стандартов и технологий при Министерстве торговли США описывает возможные риски внедрения ИИ, а также дает рекомендации по их устранению. Применение руководства носит добровольный характер, организации могут применять его, исходя из своих потребностей.
- *Руководство по вопросам использования ИИ* при принятии отрицательных решений в отношении заемщика, подготовленное Бюро по защите потребителей на финансовом рынке. В соответствии с руководством кредитор не может принять отрицательное решение в отношении заемщика без указания конкретной и обоснованной причины⁹⁸.
- *Рекомендация Службы регулирования отрасли финансовых услуг* для участников финансового рынка постоянно обновлять внутренние надзорные процедуры и документы внутреннего контроля над продуктами и услугами на основе технологий ИИ⁹⁹.
- *Руководство Федеральной торговой комиссии*, которое указывает на недопустимость недобросовестной или вводящей в заблуждение рекламной информации об ИИ¹⁰⁰.

⁹⁶ [Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence](#). The President Executive Order 14110 of October 30, 2023.

⁹⁷ [Blueprint for an AI Bill of Rights](#) (October 2022).

⁹⁸ [CFPB Issues Guidance on Credit Denials by Lenders Using Artificial Intelligence](#). Consumer Financial Protection Bureau (September 2023).

⁹⁹ [Artificial Intelligence \(AI\) in the Securities Industry](#). FINRA (June 2020).

¹⁰⁰ [Keep your AI claims in check](#). Federal Trade Commission (February 2023).

О борьбе с дискриминацией и предвзятостью в автоматизированных системах

В апреле 2023 года Министерство юстиции, Федеральная торговая комиссия, Комиссия по равным возможностям трудоустройства, Бюро по защите потребителей на финансовом рынке США опубликовали совместное заявление¹, в котором высказали обеспокоенность потенциальными рисками использования автоматизированных систем и указали на необходимость соблюдать требования законодательства о запрете всех видов дискриминации и недобросовестных методов конкуренции и бизнес-практик, независимо от применяемых технологий.

Бюро по защите потребителей на финансовом рынке предприняло следующие шаги²:

- *Циркуляр по защите прав потребителей*: в соответствии с требованиями закона о недискриминации³ кредитные организации не освобождаются от обязанности разъяснять причины отклонения заявки на кредит, даже если такие решения были приняты на основе кредитных моделей со сложными алгоритмами (black box algorithms).
- *Разъяснение о применении Закона о финансовой защите потребителей в отношении цифрового маркетинга*: организации или специалисты в области цифрового маркетинга признаются поставщиками соответствующих услуг. Результат их действий, включая применение алгоритмов для определения целевой аудитории, не должен приводить к ущербу для потребителей. В противном случае к ним будут применены меры за нарушение положений закона о защите прав потребителей.
- *Заявление о недопустимости неправомерного использования технологии ИИ на потребительских рынках*. Запрет, в частности, касается применения ИИ для сокрытия важных характеристик продукта или услуги, отказа от ответственности или использования пробелов в знаниях потребителей о технологии.

Таким образом, инновационные технологии, как и бизнес-процессы, которые они заменяют, не должны приводить к нарушениям требований законодательства по защите прав потребителей.

¹ [Joint Statement on Enforcement Efforts against Discrimination and Bias in Automated Systems](#). Department of Justice's Civil Rights Division, Federal Trade Commission, Equal Employment Opportunity Commission, Consumer Financial Protection Bureau.

² Consumer Financial Protection Circular 2022-03: [Adverse action notification requirements in connection with credit decision based on complex algorithms](#).

³ [Equal Credit Opportunity Act](#).

Канада

В июне 2022 года в парламенте Канады был разработан межотраслевой законопроект C-27, основанный на трех инициативах по защите частной жизни и информации, одна из которых – законопроект об ИИ и данных (AIDA)¹⁰¹. AIDA содержит требования к разработчикам и поставщикам продуктов ИИ по обезличиванию данных и ведению учета и устанавливает более строгие правила к высокоэффективным системам ИИ в части управления рисками.

В Канаде активно применяются механизмы мягкого регулирования применения ИИ на финансовом рынке. В апреле 2023 года Офис суперинтенданта финансовых учреждений (OSFI) в партнерстве с Институтом глобальных рисков (GRI) основал Форум финансовой индустрии по искусственному интеллекту (FIFAI), куда вошли представители регуляторов, финансовых организаций и исследовательских центров¹⁰². Форум опубликовал доклад о принципах безопасного и ответственного применения ИИ¹⁰³, к которым, в частности, относятся объяснимость алгоритмов, качественное управление данными, адекватная структура управления и соответствие этическим принципам.

¹⁰¹ [Bill C-27: An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act](#). Законопроект представлен в Палату общин Канады в июне 2022 года и на данный момент находится на этапе рассмотрения.

¹⁰² [Financial Industry Forum on Artificial Intelligence: A Canadian Perspective on Responsible AI](#). Government of Canada (June 2023).

¹⁰³ [A Canadian Perspective on Responsible AI](#). Financial Industry Forum on Artificial Intelligence (April 2023).

Китай

В Китае принимаются законы, направленные на регулирование отдельных случаев использования ИИ. В марте 2022 года вступили в силу *Правила алгоритмического управления информационными службами Интернета (IISARM¹⁰⁴)*, которые распространяются на все алгоритмические технологии, применяемые для разработки рекомендаций. Согласно требованиям данного документа, организациям запрещено использовать алгоритмы, анализирующие привычки, предпочтения, транзакции и иные данные потребителей для предложения необоснованно дифференцированных услуг.

В январе 2023 года вступили в силу *Правила регулирования дипфейков¹⁰⁵*. Документ запрещает публикацию контента, созданного ИИ, без специальной маркировки. Правила касаются не только дипфейков, но и других продуктов глубокого синтеза или генеративного ИИ.

В августе 2023 года в Китае вступило в силу первое в мире руководство по регулированию *генеративного ИИ*. Правила запрещают генерацию контента, связанного с подрывом государственной власти, пропагандой терроризма, разжиганием межнациональной вражды, распространением недостоверной информации и так далее. Регулирование предъясвляет к поставщикам решений генеративного ИИ ряд требований:

- 1) обучающие данные должны иметь законное происхождение;
- 2) ИИ не должен нарушать права лиц на объекты интеллектуальной собственности;
- 3) при использовании персональных данных для обучения ИИ необходимо согласие лица на их обработку;
- 4) поставщики ИИ должны принимать запросы физических лиц об исправлении недостоверной информации;
- 5) контент, созданный ИИ, должен быть маркирован¹⁰⁶.

3. Юрисдикции, использующие элементы стимулирующего подхода к регулированию искусственного интеллекта

Великобритания

Департамент науки, инноваций и технологий Великобритании в марте 2023 года опубликовал доклад «Инновационный подход к регулированию искусственного интеллекта», где предлагает регулировать использование ИИ, а не саму технологию и отказаться от разработки отдельных правил или уровней риска для целых секторов и технологий – в этой части доклад противопоставляется подходу, предлагаемому в проекте Регламента ЕС.

Ранее, в октябре 2022 года, Банк Англии выпустил доклад для общественных консультаций, где описаны выгоды и риски от использования ИИ на финансовом рынке, а также рассматриваются вопросы применимости действующего законодательства при использовании ИИ на финансовом рынке. Доклад не содержит конкретного комплекса мер или предложений к действию и направлен прежде всего на обсуждение вопроса о достаточности действующего регулирования или целесообразности введения дополнительного регулирования ИИ на финансовом рынке¹⁰⁷.

В Великобритании опубликовано значительное количество руководств, которые затрагивают вопросы этичного использования ИИ в разных ситуациях и сферах: в публичном секторе,

¹⁰⁴ [Internet Information Service Algorithmic Recommendation Management Provision](#). Cyberspace Administration of China. Закон разработан на основе законов Китая о кибербезопасности, защите данных и личной информации и дополняет их, а также регламент по администрированию информационных интернет-сервисов и иные аналогичные акты (Cybersecurity Law of the PRC, The PRC Data Security Law, the Personal Information Protection Law of the PRC, The Measures on the Administration of Internet Information Services).

¹⁰⁵ [Положения об управлении углубленным синтезом информационных услуг Интернета](#). Администрация киберпространства КНР (2022).

¹⁰⁶ [China's New AI Regulations](#). Latham & Watkins (August 2023).

¹⁰⁷ [DP5/22 – Artificial Intelligence and Machine Learning](#). Bank of England (October 2022).

при закупках, при автоматизированном принятии решений и так далее¹⁰⁸. Среди этических принципов при автоматизированном принятии решений указаны тестирование, определенность в вопросе назначения ответственного лица, безопасной обработки данных, раскрытия информации и соблюдение требований законодательства¹⁰⁹.

Сингапур

Монетарное управление Сингапура (MAS) в 2018 году разработало *Принципы справедливости, этики, подотчетности и прозрачности (FEAT)*, которые были обновлены в 2019 и 2020 годах в целях имплементации предложений Комиссии по защите персональных данных Сингапура¹¹⁰. В частности, нововведения затрагивают следующие вопросы:

- внутреннее управление и меры;
- модели ИИ-решений, в том числе направленные на определение приемлемого риска и обстоятельств применения методик, основанных на различных форматах обучения с разной степенью привлечения человека;
- операционное управление, включая формирование отчетности о данных и сведение к минимуму возможных погрешностей;
- управление взаимоотношениями с клиентами, включая раскрытие информации, прозрачность и объяснимость.

Важный документ, который задает ориентиры политики Сингапура в вопросах права и этики ИИ, – *Рамочная модель регулирования ИИ (The Model AI Governance Framework¹¹¹)*. В части обеспечения добросовестности выделены необходимость защиты от систематического ущемления интересов потребителей, регулярная проверка моделей на корректность работы и исключение систематических ошибок.

В 2022 году в Сингапуре была создана первая в мире система тестирования управления ИИ (AI Verify), в рамках которой разработчики ИИ могут самостоятельно верифицировать соответствие *Рамочной модели регулирования ИИ*. Для этих целей было разработано *Руководство по внедрению и самостоятельной оценке для организаций (Implementation and Self-Assessment Guide for Organisations, ISAGO¹¹²)*.

6.2. Подходы к регулированию искусственного интеллекта на российском финансовом рынке

На данный момент Банк России считает целесообразным поддерживать создание условий, направленных на стимулирование развития ИИ на финансовом рынке с учетом риск-ориентированного принципа регулирования ИИ.

В целях содействия развитию технологии ИИ Банк России продолжит принимать участие в реализации следующих направлений:

1. Развитие регулирования оборота **обезличенных персональных данных**. В целях обеспечения благоприятных правовых условий для сбора, хранения и обработки данных с использованием новых технологий при участии Банка России был разработан законопроект, направленный на установление порядка обезличивания персональных данных для

¹⁰⁸ [Data ethics and AI guidance landscape](#). Central Digital and Data Office, Cabinet Office, and Office for Artificial Intelligence (May 2021).

¹⁰⁹ [Guidance Ethics, Transparency and Accountability Framework for Automated Decision-Making](#). Central Digital and Data Office, Cabinet Office, and Office for Artificial Intelligence (May 2021).

¹¹⁰ [Principles to Promote Fairness, Ethics, Accountability and Transparency \(FEAT\) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector Monetary Authority of Singapore](#). Monetary Authority of Singapore (November 2018).

¹¹¹ [Model AI Governance Framework](#). PDPC (January 2019 – 1st issue, January 2020).

¹¹² [Companion to the Model AI Governance Framework – Implementation and Self-Assessment Guide for Organizations](#). IMDA, PDPC, WEF (January 2020).

операторов персональных данных, не являющихся государственными органами¹¹³. Принятие законопроекта позволит минимизировать риски при использовании обезличенных данных для обучения моделей ИИ. Законопроект предусматривает совершенствование правового режима согласий на обработку персональных данных, что упростит получение законного основания на использование данных для разработки решений в области ИИ.

2. Возможность использования разработчиками ИИ **сторонней инфраструктуры обработки данных**. Вопрос особенно актуален для обучения моделей ИИ, требующих значительных информационных и аппаратных ресурсов. Использование облачных решений может снизить издержки участников рынка. Так, при участии Банка России разработан законопроект¹¹⁴, направленный на предоставление возможности финансовым организациям передавать для обработки информацию, составляющую банковскую тайну или иной вид тайны, при аутсорсинге функций, связанных с использованием облачных услуг.
3. Формирование подходов к распределению ответственности между разработчиком технологии ИИ и организацией-пользователем за вред, причиненный в результате применения ИИ, а также к иным вопросам, требующим дополнительной проработки на уровне законодательства: определение правового режима объектов, созданных ИИ, в том числе с позиций законодательства об интеллектуальной собственности (что особенно актуально для генеративных моделей ИИ), определение правового режима моделей ИИ как результата интеллектуальной деятельности.

Проработка указанных вопросов позволит создать прозрачную правовую среду для разработчиков ИИ, что придаст импульс созданию и внедрению новых решений на основе ИИ.

При этом важно отметить, что технологии ИИ носят универсальный характер, их применение не ограничивается конкретной сферой, а решения на основе ИИ существенно влияют на все отрасли экономики, социальной сферы и государственного управления. В связи с этим угрозы, которые несут риски внедрения ИИ, могут повлиять как на стабильность финансового рынка, так и на другие отрасли экономики и общественные отношения. Одновременно с этим избыточное регулирование может не только ограничить возможности разработки и внедрения продуктов и сервисов на финансовом рынке, но и оказать негативное воздействие на развитие технологий ИИ в других отраслях. Не все области применения ИИ порождают риски, требующие вмешательства регулятора.

В связи с этим регулирование ИИ должно носить пропорциональный характер, в основе которого лежит **применение риск-ориентированного принципа**, который подразумевает зависимость формата регуляторных требований от объема и вероятности реализации риска использования технологии в сравнении с потенциальным положительным эффектом от ее внедрения. Использование риск-ориентированного подхода позволит соблюсти баланс между созданием условий для развития ИИ и управлением рисками, требующими внимания со стороны регулятора.

В рамках реализации **Национальной стратегии развития ИИ до 2030 года** при участии Банка России в 2020 году принята **Концепция развития регулирования отношений в сфере технологий ИИ и робототехники до 2024 года**. Согласно данной концепции, в основе регулирования ИИ в России должны лежать применение **риск-ориентированного принципа**, использование инструментов сорегулирования¹¹⁵ и саморегулирования, а также формирование этических правил разработки и применения ИИ.

На данный момент Банк России поддерживает целесообразность использования риск-ориентированного подхода к регулированию ИИ.

¹¹³ Законопроект № 992331-7 принят 26.01.2021 Государственной Думой в первом чтении.

¹¹⁴ Законопроект № 404786-8 в настоящее время рассматривается Государственной Думой.

¹¹⁵ Во взаимодействии нескольких регуляторов.

Данный подход предполагает сочетание инструментов так называемого *мягкого регулирования, саморегулирования и точечного нормативного регулирования* пропорционально объему и вероятности реализации рисков внедрения технологий ИИ.

В рамках риск-ориентированного принципа регулирования ИИ представляется возможным выделить следующие категории рисков:

- **Риски, не требующие вмешательства регуляторов.**

При регулировании рисков, возникающих в связи с внедрением ИИ на финансовом рынке, необходимо учитывать, что не все риски требуют вмешательства регулятора. При этом использование ИИ для внутренних процессов в компании может стать предметом интереса, если такие процессы уже регулируются действующим законодательством, а использование ИИ для таких процессов может помешать достижению целей действующего законодательства.

- **Риски, которые могут быть ограничены за счет саморегулирования и рекомендаций.**

Другая часть рисков может быть проработана с помощью инструментов саморегулирования и мягкого регулирования. При таком подходе объединения организаций или регуляторы могут выпускать не имеющие обязывающего характера руководства, которые содержат этические принципы разработки и внедрения ИИ, а также рекомендации по иным вопросам (например, рекомендации относительно использования ИИ, чтобы при этом не выходить за рамки действующего законодательства).

Применение механизмов проработки рисков использования ИИ на основе саморегулирования – общемировая тенденция. Одним из важных инструментов управления рисками является добровольное принятие участниками рынка определенных **этических принципов** (как правило, в виде кодексов, хартий) использования ИИ. Как показывает мировая практика, инициатива создания подобных объединений может исходить как от регуляторов, так и от участников рынка. Такой подход позволяет без создания дополнительных регуляторных барьеров прорабатывать случаи использования технологии с низким риском, когда управление таким риском не требует со стороны регулятора нормативного закрепления правил, обязательных для исполнения участниками рынка.

Значительная часть документов этического использования ИИ предусматривает приверженность участников рынка принципам информационной открытости о технологиях ИИ и потенциале их применения; прозрачности процесса принятия решений на основе ИИ; законности при разработке, внедрении и использовании ИИ; безопасности при использовании ИИ; неиспользования алгоритмов ИИ, приводящих к дискриминации определенных социальных групп.

В России уже есть примеры добровольного объединения участников рынка для формирования и применения этических принципов (в том числе использования качественных наборов данных на этапе обучения моделей, соблюдения безопасности при передаче данных и другое) при создании, внедрении и использовании технологий ИИ¹¹⁶. Использование данных инструментов позволяет не только осуществлять добровольную и гибкую проработку технических рисков, но и в рамках совместной деятельности обобщать лучшие практики управления такими рисками и делать их доступными для других участников рынка.

Обмен лучшими практиками в сфере ИИ между участниками рынка является важным элементом проработки рисков, связанных с использованием технологии. В рамках такой работы предлагается рассмотреть вопрос о подготовке участниками рынка совместно с Банком России **Кодекса этики применения ИИ на финансовом рынке**, учитывающего специфику применения ИИ при оказании финансовых услуг. Добровольное соблюдение финансовыми организациями такого Кодекса может содействовать повышению доверия к оказанию финансовых услуг с применением ИИ и защите интересов клиентов финансовых организаций.

¹¹⁶ Такие этические нормы закреплены в Кодексе этики Альянса в сфере ИИ, а также Кодексе этики использования данных. Эти кодексы объединяют 209 организаций из различных отраслей и сфер деятельности (операторы сотовой связи, ИТ-компании, научные организации и так далее), в том числе участников финансового рынка – банки, управляющие компании и другие.

По инициативе финансовых организаций возможно **привлечение Банка России для участия в наблюдении за проектами по внедрению ИИ и анализе связанных с ними рисков.** Возможность проводить предварительные консультации с регулятором и тестировать новые решения и продукты с его участием, в том числе в рамках регулятивной «песочницы» Банка России, позволит финансовым организациям и регулятору полнее оценить возникающие риски и предусмотреть механизмы, необходимые для их ограничения.

- **Риски, требующие внимания регуляторов.**

Действующее российское регулирование финансового рынка в ряде случаев уже устанавливает требования к применению различных моделей, в том числе основанных на технологиях традиционного ИИ, в деятельности финансовых организаций. Например, в банковской сфере установлены требования к валидации банковских моделей оценки рисков¹¹⁷ и к управлению кредитными организациями операционными, в том числе модельными рисками¹¹⁸. В инвестиционной сфере предусмотрена специальная процедура аккредитации программ для ЭВМ, посредством которых осуществляется предоставление индивидуальных инвестиционных рекомендаций¹¹⁹.

При этом Банк России отмечает **важность самостоятельного управления финансовыми организациями рисками**, связанными с применением ИИ, и выстраивания в финансовых организациях, применяющих ИИ, системы управления такими рисками, учитывающей характер и масштаб таких рисков, а также особенности применяемых программ на основе ИИ.

Вместе с тем широкое распространение использования генеративного ИИ и расширение сфер использования технологии ИИ в целом на финансовом рынке могут быть сопряжены с появлением новых, специфических для данной технологии рисков, требующих дополнительного анализа и оценки. Выявление факторов, значимых для определения необходимости регуляторного вмешательства, будет осуществляться в рамках систематического мониторинга Банка России. Необходимость разработки отдельных подходов и методов регулирования этих рисков целесообразно оценивать по результатам всестороннего обсуждения с участниками рынка с учетом риск-факторов.

В таком случае при введении новых требований целесообразно проводить тщательную оценку регуляторного воздействия таких требований, в том числе с возможностью использования **механизма пилотирования цифровых инноваций** (в том числе в рамках экспериментальных правовых режимов), а также оценивать их пропорциональность уровню риска, возникающему при внедрении и эксплуатации технологии ИИ.

При формировании необходимых требований Банк России будет использовать **принципы риск-ориентированного подхода**, позволяющего финансовым организациям внедрять новые технологии и развивать ИИ, при этом обеспечивая необходимый уровень системы управления рисками.

Банк России в целом придерживается технологически нейтрального подхода к регулированию, позволяющего финансовым организациям внедрять новые технологии и развивать ИИ (к примеру, требований, аналогичных текущим требованиям к точности банковских моделей оценки рисков, не зависящим от применения ИИ или иных алгоритмов при проведении расчетов, или ограничений, аналогичных установленному сегодня запрету использовать дискриминирующие детерминанты при определении стоимости полиса ОСАГО независимо от технологии скоринга), и **на данный момент не видит необходимости в оперативной разработке отдельного регулирования использования технологии ИИ финансовыми организациями.**

¹¹⁷ Положение Банка России от 24.08.2020 № 730-П.

¹¹⁸ Положение Банка России от 08.04.2020 № 716-П.

¹¹⁹ Указание Банка России от 27.11.2018 № 4980-У.

Внедрение специальных требований, ориентированных исключительно на ИИ (к примеру, сертификации алгоритмов, требований к их архитектуре или размеру обучающих выборок), **может быть рассмотрено в отдельных случаях после дополнительного анализа и изучения**, так как указанные требования менее адаптивны и их пересмотр может не успевать за развитием технологий ИИ.

ВОПРОСЫ ДЛЯ ОБЩЕСТВЕННОГО ОБСУЖДЕНИЯ

Вопрос к разделу 2 «Глобальные тренды развития искусственного интеллекта»

1. Какие дополнительные тренды развития ИИ на финансовом рынке вы можете выделить?

Вопросы к разделу 3 «Применение технологий искусственного интеллекта на финансовом рынке»

2. Учитываете ли вы применение ИИ в стратегических документах вашей организации (общие документы, бизнес-стратегии, стратегии цифровой трансформации и так далее)?
3. Оцените, пожалуйста, в какой мере применение ИИ влияет на внутренние процессы в вашей организации (по шкале от 0 до 10, где 0 – неприменение, а 10 – автономное использование), в частности:
 - управление рисками;
 - ПОД/ФТ, антифрод;
 - разработка ПО и так далее;
 - взаимодействие с клиентами;
 - иное.
4. Считаете ли вы, что внедрение ИИ приведет к уменьшению стоимости финансовых продуктов и услуг для конечного потребителя?
5. Каковы ваши прогнозы в части развития ИИ на финансовом рынке? Каково ваше видение трансформации финансового рынка/сектора/продуктов под влиянием ИИ на горизонте 3–5 лет?

Вопросы к разделу 4 «Условия развития искусственного интеллекта на российском финансовом рынке»

6. Согласны ли вы с актуальностью и важностью обозначенных в докладе условий развития ИИ?
7. Какие дополнительные условия развития ИИ вы можете выделить?
8. Использование ИИ средними и небольшими участниками: что может способствовать массовому внедрению таких технологических решений?
9. Считаете ли вы целесообразным формирование консорциумов участников для совместной разработки ИИ?
10. Какие технологические решения, создаваемые на базе консорциума участников рынка, были бы полезны?
11. Какие виды данных в области финансового рынка от государственных органов власти необходимы для дальнейшего развития ИИ-решений?

Вопросы к разделу 5 «Риски искусственного интеллекта на финансовом рынке»

12. Согласны ли вы с обозначенными в докладе рисками, связанными с внедрением технологий ИИ?
13. Какие риски использования ИИ на финансовом рынке вы считаете наиболее весомыми с точки зрения вероятности, а также степени потенциального ущерба от их реализации?
14. Какие дополнительные риски использования ИИ на финансовом рынке вы могли бы выделить?
15. Какие качественные и количественные показатели можно использовать для оценки значимости рисков использования ИИ на финансовом рынке?

16. Какие технологии ИИ вы сейчас не используете, поскольку потенциальные риски превышают потенциальную пользу?
17. Раскрываете ли вы информацию клиентам о том, что они взаимодействуют с ИИ? Если да, на каком этапе?
18. Есть ли у ваших клиентов возможность отказаться от взаимодействия с ИИ? Если да, уточните, пожалуйста, процедуру.

Вопросы к разделу 6 «Подходы к регулированию искусственного интеллекта»

19. Согласны ли вы с предложенным подходом к регулированию ИИ на финансовом рынке?
20. Согласны ли вы с перечисленными в докладе направлениями для содействия развитию технологии ИИ, в обсуждении которых Банк России планирует продолжать принимать участие? Какие еще направления для содействия развитию технологии ИИ (не перечисленные в докладе) вы можете обозначить?
21. Являются ли этические принципы, руководства и необязательные стандарты необходимым элементом регулирования отдельных рисков ИИ?
22. Считаете ли вы целесообразной проработку вопроса создания Кодекса этики ИИ для участников финансового рынка, учитывающего специфику применения ИИ при оказании финансовых услуг?
23. Могут ли, по вашему мнению, подходы к управлению операционными рисками, установленные для кредитных организаций, быть рекомендованы к применению с учетом характера, масштаба и специфики деятельности иным финансовым организациям, влияние ИИ на оказание услуг клиентам которых или на финансовую устойчивость которых, по их оценке, является существенным?
24. Целесообразно ли создание централизованного органа внутри организации по контролю за внедрением ИИ и его применением в бизнес-процессах, этики использования, контроля и оценки рисков, последствий их реализации?
25. Приведите примеры, когда действующее законодательство не учитывает риски, возникающие в связи с использованием ИИ на финансовом рынке. В каких случаях и сферах может потребоваться адаптация действующего законодательства или принятие новых нормативных актов?

ГЛОССАРИЙ

Адаптивный ИИ – тип ИИ, в рамках которого модель может обучаться и динамически адаптироваться к изменениям в окружающей среде за счет сбора обратной связи в режиме реального времени

Генеративный ИИ – тип ИИ, который может создавать новый контент (текст, изображение, аудио и видео) в ответ на запросы пользователя

Глубокое обучение (Deep Learning) – вид машинного обучения на основе искусственных нейронных сетей (Artificial Neural Networks) – моделей ИИ, которые представляют собой систему соединенных и взаимодействующих между собой логических вычислительных единиц (искусственных нейронов). Целью создания искусственных нейронных сетей была попытка имитировать процесс человеческого мышления, модель строится по тем же принципам, что и нервная система человека

Задача выявления аномалий – выявление аномальных значений в выборке стандартных данных

Задача кластеризации – распределение по группам отдельных экземпляров данных на основании наличия у них схожих признаков

Задача регрессии – прогнозирование на основе данных, обладающих различными признаками

Искусственный интеллект (ИИ) – комплекс технологических решений, позволяющий имитировать когнитивные функции человека и получать при выполнении конкретных задач результаты, сопоставимые с результатами его интеллектуальной деятельности. Комплекс технологических решений включает в себя программное обеспечение, информационно-коммуникационную инфраструктуру, процессы и сервисы по обработке данных и поиску решений

Компьютерное зрение (Computer Vision) – технология, направленная на извлечение информации из визуальных данных и принятие решений на основе этой информации

Машинное обучение – класс алгоритмов ИИ, направленный на поиск решения конкретной задачи посредством обучения моделей на выборке решений множества сходных задач

Молчаливый сговор – форма антиконкурентной координации, которая может быть достигнута без какой-либо необходимости в явном соглашении, но которое участники рынка могут поддерживать, признавая свою взаимозависимость

Мультимодальное обучение (Multimodal Learning) – обучение, посредством которого модели ИИ могут предоставлять пользователю ответ с использованием различных форматов информации – визуального, звукового, тактильного, текстового

Обработка естественного языка (Natural Language Processing, NLP) – технология ИИ, которая позволяет компьютерам: 1) понимать (Natural language understanding, NLU) и 2) генерировать (Natural language generating, NLG) текстовую информацию на естественных языках

Обучение без учителя (Unsupervised Learning) – вид машинного обучения, при котором модель не получает заранее размеченных данных, а выявляет закономерности между ними самостоятельно. Типичным примером обучения без учителя является обнаружение аномалий (например, мониторинг транзакций для ПОД/ФТ)

Обучение с подкреплением (Reinforcement Learning) – обучение модели во взаимодействии со средой и за счет оптимизации ее поведения на основе получаемого за различные действия вознаграждения

Обучение с учителем (Supervised Learning) – вид машинного обучения, при котором обучение модели происходит на входных данных, предварительно размеченных для решения конкретной задачи

Обучение Low-code/No-code – подход к разработке программного обеспечения, который позволяет создавать приложения либо с минимальным написанием кода, либо без него

Открытые API – программные интерфейсы, публикуемые организациями в соответствии с требованиями Банка России для обеспечения возможности цифрового обмена данными с поставщиками услуг (с согласия клиента) и клиентами в рамках организации и предоставления финансовых услуг

Разметка данных (Data Labeling) – процесс добавления меток в сырые данные, нацеленный на то, чтобы показать модели целевые ответы, которые она должна предсказывать

Распознавание и синтез речи (Automatic Speech Recognition & Text-To-Speech) – комплекс технологий, направленный на преобразование речи человека в цифровую информацию и имитацию речи

Сильный ИИ (Artificial General Intelligence, AGI – ИИ общего назначения) – гипотетическая концепция системы ИИ, которая способна к решению любого класса задач, которые могут быть выполнены человеком. Среди возможных свойств сильного ИИ выделяют способность к самообучению, автономному принятию решений и самостоятельной постановке задач, построению логических умозаключений, стратегическому планированию и действиям в условиях значительной неопределенности

Синтетические данные – искусственные данные, сгенерированные специально обученной моделью ИИ, которые имитируют характеристики реальных данных и могут быть использованы для обучения других моделей ИИ

Слабый ИИ (Artificial Narrow Intelligence, ANI – ИИ узкого назначения) определяется как набор алгоритмов, ориентированный на решение узкоспециализированных, прикладных задач

Суперсильный ИИ (Artificial Superintelligence, ASI – супер-ИИ) – концепция, выделяемая некоторыми исследователями, которая описывает возможность создания в отдаленной перспективе систем ИИ, которые будут существенно превосходить разум человека

Традиционный ИИ – тип ИИ, имитирующий умственную деятельность и способный делать прогнозы и выводы на основе анализа. В финансовой отрасли используется для кредитного скоринга, борьбы с мошенничеством, управления рисками и других задач

Федеративное обучение (Federated Learning) – способ обучения моделей ИИ на основе нескольких независимых наборов данных (источники данных могут быть разных видов)

Этика ИИ – совокупность принципов и методов их реализации для создания и ответственного использования технологии ИИ в целях предотвращения рисков, связанных с нарушениями в области морали и нравственности

Явный сговор – антиконкурентное поведение, которое подтверждается явными соглашениями (письменными или устными)

RegTech (Regulatory Technology) – технологии, используемые для упрощения выполнения требований регуляторов

СПИСОК ЛИТЕРАТУРЫ

1. Администрация киберпространства КНР (2022). [Положения об управлении углубленным синтезом информационных услуг Интернета](#)
2. Ассоциация ФинТех (23.08.2023). [Искусственный интеллект – основа для создания финансовых услуг нового поколения](#)
3. Ассоциация ФинТех (05.10.2023). [Применение технологий искусственного интеллекта на финансовом рынке \(расширенная версия\)](#)
4. ЮНЕСКО (2012). [Рекомендации об этических аспектах искусственного интеллекта](#)
5. Arthur D. Little (October 2023). [Generative Artificial Intelligence: toward a new civilization?](#)
6. Bank of England (October 2022). [DP5/22 – Artificial Intelligence and Machine Learning](#)
7. Bank of Italy (October 2022). Emilia Bonaccorsi di Patti, Filippo Calabresi and etc. [No. 721 – Artificial intelligence in credit scoring: an analysis of some experiences in the Italian financial system](#)
8. BIS (August 2021). [Artificial intelligence and machine learning in financial services](#)
9. CFA (2020). [Artificial Intelligence in Asset Management](#)
10. Deutsche Bank (June 2021). [Unleashing the potential of AI in securities services](#)
11. EBA (2020). [Guidelines on loan origination and monitoring](#)
12. ESMA (01.02.2023). [TRV Risk Analysis Artificial intelligence in EU securities markets](#)
13. European Parliament (2020). [The ethics of artificial intelligence: Issues and initiatives](#)
14. Gartner (17.10.2022). Gartner – Top Strategic Technology Trends 2023
15. Harvard (2017). [History of Artificial Intelligence](#)
16. IMF (October 2021). [Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance](#)
17. IMF (August 2023). [Generative Artificial Intelligence in Finance: Risk Considerations](#)
18. IOSCO (2020). [The Use of Artificial Intelligence and Machine Learning by market intermediaries and asset managers](#)
19. KPMG (April 2023) [Generative AI models – the risks and potential rewards in business](#)
20. McKinsey (2023). [The economic potential of generative AI: The next productivity frontier](#)
21. McKinsey (2023). [What's the future of generative AI? An early view in 15 charts](#)
22. McKinsey (2022). [An executive's guide to AI](#)
23. NASA (April 2021). [NASA Framework for the Ethical Use of Artificial Intelligence \(AI\)](#)
24. OECD (May 2019). [Recommendation of the Council on Artificial Intelligence](#)
25. OECD (February 2023). [A blueprint for building national compute capacity for artificial intelligence](#)
26. OECD (18.09.2023). [Initial policy considerations for generative artificial intelligence](#)
27. Stanford University (2023). [Artificial Intelligence Index Report 2023](#)
28. UN (July 2023). [Artificial Intelligence: Opportunities and Risks for International Peace and Security – Security Council, 9381st Meeting](#)