ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ПОД ЗАЩИТОЙ

РОЖДЕНИЕ РЫНКА ЗАЩИТЫ ИИ: ТЕКУЩЕЕ СОСТОЯНИЕ И ПЕРСПЕКТИВЫ

Сентябрь 2026 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Искусственный интеллект (ИИ) стремительно трансформируется в новую базовую платформу для всего ИТ-рынка. С развитием ИИ закономерно усиливается и его значимость как **объекта защиты**. В отличие от традиционных ИТ-систем ИИ-решения опираются на данные, обучаемые модели и зачастую ограниченно интерпретируемую логику принятия решений. Это формирует новый класс угроз — от атак на модели и данные до злоупотреблений генеративными системами, утечек через ИИ-интерфейсы, а также до неконтролируемого поведения ИИ-агентов и приложений. В результате **кибербезопасность ИИ перестает быть нишевой темой** и становится важным направлением развития рынка решений информационной безопасности (ИБ).

Осознавая масштаб и скорость этих изменений, в рамках совместного исследования Группа компаний Б1 (Б1), ГК «Солар», Ассоциация ФинТех (АФТ), HiveTrace поставили перед собой задачу понять, **как развивается рынок средств защиты ИИ в России и мире**. Мы опросили порядка 60 организаций из различных отраслей, включая финансы, ритейл, промышленность, телеком и государственный сектор в России.

Важно отметить, что в данном исследовании рассматривается защита ИИ-систем от злоумышленников, кибератак и сбоев, способных нарушить работу ИТ-контура и бизнес-процессы компании, а не направление безопасного и этичного использования ИИ, нацеленное на предотвращение рисков, угроз и преднамеренного или непреднамеренного вреда человечеству со стороны искусственного интеллекта.

Вторая тематика важна, но является отдельной дисциплиной и не находится в фокусе данного документа. Вместе с тем защита ИИ от киберугроз представляет собой важную предпосылку общей безопасности ИИ.

Результаты исследования показывают, что глобальный рынок решений и услуг в сфере кибербезопасности ИИ стремительно развивается. Появляются **новые разработчики**, а существующие ведущие ИБ-поставщики **создают платформенные решения и встраивают функционал защиты ИИ** в свои продукты и услуги.

Рынок ИБ для ИИ в РФ находится на начальной стадии формирования. Многие компании только осознают специфику рисков и приступают к адаптации существующих инструментов под новые задачи. В то же время уже прослеживаются ключевые направления развития: формирование специализированных решений для защиты моделей и данных, усиление контроля за использованием генеративного ИИ, рост внимания к вопросам управления рисками и комплаенса.

Несмотря на раннюю стадию развития, можно ожидать, что защита ИИ превратится в важное направление внутри рынка средств и услуг информационной безопасности. Для компаний это означает необходимость уже сегодня закладывать основы безопасного использования ИИ, определять приоритеты инвестиций и выстраивать стратегию защиты с учетом новых технологических реалий.

Б1, Солар, АФТ, HiveTrace

ИИ ЯВЛЯЕТСЯ КЛЮЧЕВОЙ ПЛАТФОРМОЙ РАЗВИТИЯ ИТ ВО ВСЕМ МИРЕ

~1,76 трлн

долл. США составили общие траты и инвестиции на ИИ и инфраструктуру для ИИ в 2025 году

88%

организаций в мире уже используют ИИ хотя бы в одной из бизнес-функций

>26%

всех ИТ-трат бизнеса будет связано с ИИ (ПО, услуги, инфраструктура) к 2029 г.

1+ млрд

ИИ-агентов будет активно работать в мире к 2029 г. **40+%** всего прикладного ПО будет включать ИИ-агентов к 2026–2027 гг.

Источники: IDC, Gartner, McKinsey

ИИ СТАНОВИТСЯ НОВОЙ КРИТИЧЕСКИ ВАЖНОЙ ИНФРАСТРУКТУРОЙ, А ЗАЩИТА ИИ — ОТДЕЛЬНЫМ ВАЖНЫМ КЛАССОМ ТЕХНОЛОГИЙ ИБ

Искусственный интеллект стремительно превращается в ключевой элемент цифровой инфраструктуры аналогично тому, как ранее критически важными стали облачные платформы, мобильные экосистемы, ПО для автоматизации бизнес-процессов и корпоративные сети.

Уже в ближайшей перспективе более четверти всех ИТ-расходов будут так или иначе связаны с технологиями искусственного интеллекта. Аналитические агентства прогнозируют, что к 2030 году ИИ будет задействован практически во всех ИТ-процессах, а до 25% задач будут выполняться ИИ-системами автономно.

Компании активно ускоряют внедрение ИИ. По данным «Cisco AI Readiness Index», более 90% организаций на ряде рынков планируют развертывание ИИ-агентов в ближайшие годы. Однако готовность инфраструктуры, процессов управления и механизмов безопасности отстает от темпов внедрения ИИ.

Организации по всему миру сталкиваются с реальными угрозами ИИ-системам.

Более того, риски, связанные с безопасностью ИИ, – это серьезный барьер для масштабного внедрения генеративного ИИ (ГенИИ).

Традиционные средства ИБ уже недостаточны для защиты ИИ-моделей и ИИ-систем с дополненным поиском по базам знаний (RAG), автономных ИИ-агентов и экосистем инструментов ИИ, поскольку ключевые угрозы возникают на уровне данных, логики поведения моделей и их самостоятельных действий.

По мере масштабирования ИИ и его интеграции в бизнес-процессы и системы риски манипуляции запросами, неконтролируемой работы агентов, утечек конфиденциальных данных и компрометации баз знаний будут усиливаться.

Осознание этой проблемы ведет к формированию принципиально нового класса технологий рынка ИБ – безопасности систем искусственного интеллекта – и к появлению специализированных подходов к защите ИИ-моделей, приложений и инфраструктуры.

Источники: Gartner, Cisco

ОРГАНИЗАЦИИ ПО ВСЕМУ МИРУ СЕРЬЕЗНО РАССМАТРИВАЮТ РИСКИ, СВЯЗАННЫЕ С ИИ, И УЖЕ ВСТРЕЧАЮТСЯ С НИМИ НА ПРАКТИКЕ

31%

организаций считают себя полностью способными защищать агентские ИИ-системы

32%

организаций в мире столкнулись с атакой на ИИ-приложения с использованием промтов

46%

организаций считают риски, связанные с безопасностью, важным барьером для масштабного внедрения ГенИИ

29%

организаций уже столкнулись с атаками на инфраструктуру приложений ГенИИ

Источники: Gartner, CBInsights

РИСКИ ИБ ДЛЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ИМЕЮТ СПЕЦИФИЧЕСКИЙ ХАРАКТЕР: МАНИПУЛИРОВАНИЕ, РИСКИ ВО ВРЕМЯ РАБОТЫ, АТАКИ НА МОДЕЛИ

Манипуляция поведением модели через входные данные и инструкции

Воздействие на поведение ИИ-системы через специально сформированные запросы, скрытые инструкции или вредоносный контент

Утечки данных и секретов через запросы

Передача в ИИ-систему конфиденциальных данных, персональной информации или секретов доступа через пользовательские запросы и интеграции

Риски автономного и неконтролируемого поведения ИИ-систем

Выполнение ИИ-агентами и автономными ИИ-системами ошибочных, небезопасных или несанкционированных действий вследствие автономности, галлюцинаций или избыточных полномочий

Уязвимости моделей и решений с открытым исходным кодом, риски цепочки поставок ИИ

Компрометация ИИ-систем через уязвимости, вредоносные зависимости или недоверенные внешние модели, библиотеки и ИИ-компоненты

Неконтролируемое, теневое, нарушающее регулирование использование ИИ

Применение ИИ-систем без централизованного контроля, governance- и compliance-процедур, создающее риски нарушения требований безопасности, конфиденциальности и регулирования

Компрометация моделей, датасетов, RAG-систем и баз знаний

Внесение вредоносных изменений в модели, обучающие данные или базы знаний, приводящее к некорректной, небезопасной или управляемой злоумышленником работе ИИ-системы

Утечки данных через ответы модели

Раскрытие ИИ-системой чувствительной информации в ответах вследствие ошибок доступа, особенностей обучения или механизмов поиска и извлечения данных

Небезопасная обработка результатов модели

Автоматическое использование или исполнение ответов ИИ-модели, включая написание кода, без проверки и контроля, приводящее к выполнению вредоносных действий и компрометации систем

Атаки на модели и ИИ-инфраструктуру

Воздействие на модели и инфраструктуру ИИ с целью кражи моделей, восстановления обучающих данных, нарушения работы или обхода механизмов защиты

Использование ИИ для автоматизации кибератак и мошенничества

Применение ИИ для масштабирования фишинга, социальной инженерии, генерации вредоносного кода, deepfake-атак и других форм кибермошенничества

Источник: анализ исследовательской группы

ВОЗМОЖНОСТЬ РЕАЛИЗАЦИИ ДАННЫХ РИСКОВ ПОДТВЕРЖДАЕТСЯ КАК НА ПРАКТИКЕ, ТАК И АКАДЕМИЧЕСКИМИ ИССЛЕДОВАНИЯМИ

Манипуляция поведением модели через входные данные и инструкции

В 2023 году был продемонстрирован критический дефект в ГенИИ-ориентированном чат-боте, развернутом компанией Fullpath на сайте автосалона. На базе метода ролевого манипулирования были переопределены базовые системные инструкции ИИ (приказ соглашаться с любыми утверждениями клиента, фраза о том, что данная оферта является юридически обязывающей)

Утечки данных и секретов через запросы

В марте 2023 года руководство дивизиона крупной технологической корпорации разрешило использование ГенИИ-приложений для решения производственных задач. В течение 20 дней было зафиксировано как минимум три независимых инцидента утечки интеллектуальной собственности и секретов производства интерфейс-ввода

Риски автономного и неконтролируемого поведения ИИ-систем

Разработанный для упрощения задач программирования ИИ-агент на базе популярного ИИ ПО сумел за считанные секунды стереть всю корпоративную базу данных и резервные копии без инструкции на действие или подтверждение данного действия, выступив в роли вредоносного ПО

Уязвимости моделей и решений с открытым исходным кодом, риски цепочки поставок ИИ

В конце 2024 года компания HiddenLayer зафиксировала атаку на цепочку поставок ИИ через создание фиктивного репозитория Open-OSS/privacy-filter на Hugging Face, маскирующегося под официальный проект OpenAI. Вредоносное ПО скрытно извлекало пароли, файлы конфигурации и приватные ключи криптовалютных кошельков

Компрометация моделей, датасетов, систем ИИ с доппоиском (RAG) и баз знаний

База знаний в системах ГенИИ, дополненных поиском, – самостоятельная и практическая поверхность атаки. Исследователи добивались 90% успеха атак после добавления 5 вредоносных текстов на целевой вопрос в базу с миллионами документов, а одного «авторитетного» документа хватило для успешной манипуляции ответом системы

Утечки данных через ответы модели

Исследователи («Извлечение обучающих данных из больших языковых моделей» USENIX Security) продемонстрировали возможность извлекать из больших языковых моделей фрагменты обучающих данных – почтовые адреса, номера телефонов, уникальные строки текста и персональные данные – с помощью специально сформированных запросов

Неконтролируемое, теневое, нарушающее регулирование использование ИИ

В 2023–2024 гг. крупнейшие мировые инвестбанки столкнулись с волной «теневого» использования ГенИИ. Аналитики и трейдеры массово начали использовать публичные веб-интерфейсы, загружая туда закрытую коммерческую информацию. Крупнейшие конгломераты были вынуждены экстренно заблокировать доступ со всех корпоративных устройств под угрозой штрафов

Атаки на модели и ИИ-инфраструктуру

Группа ИБ-исследователей из Google DeepMind, ETH Zurich и других институтов продемонстрировали возможность извлечения точных параметров последнего слоя нейросети (ее веса и размерность скрытых слоев) коммерческих ИИ-сервисов, отправляя серии специальным образом сформированных запросов к публичным API-интерфейсам

Источник: анализ исследовательской группы

РАЗВИТИЕ РЫНКА И КИБЕРУГРОЗ НЕИЗБЕЖНО ПРИВЕДЕТ К СОЗДАНИЮ НОВЫХ ПОДХОДОВ И СРЕДСТВ ЗАЩИТЫ ИИ

67% руководителей в области кибербезопасности по всему миру считают, что риски, связанные с ИИ нового поколения, требуют существенных изменений в существующих подходах к кибербезопасности.

Стремительное внедрение систем искусственного интеллекта и ИИ-агентов, которые начали получать доступ к внутренним системам и данным компаний, формирует новый класс киберрисков и существенно расширяет поверхность атак.

В ответ на это на глобальном рынке появляются специализированные средства защиты ИИ. Среди их: средства защиты ИИ-приложений; шлюзы безопасного доступа к ИИ и защита от Теневого ИИ; платформы безопасности и управления ИИ-агентами; решения для управления безопасностью ИИ-артефактов и моделей; решения защиты ИИ-систем дополненных поиском по базам знаний (RAG) и ИИ-данных; ПО для наблюдаемости и мониторинга ИИ; решения управления идентификацией, правами доступа и политиками безопасности для ИИ-агентов и ИИ-сервисов; платформы управления ИИ-рисками; решения автоматизированного тестирования моделей и ИИ-приложений; приложения для идентификации сгенерированного ИИ-контента.

По динамике расширения портфелей разработчиков видно, что статических проверок безопасности становится недостаточно. Тренд развивается в сторону мониторинга ИИ-систем и ИИ-агентов, трассировки действий и запросов, применения политик безопасности во время выполнения, непрерывного контроля и выявления угроз.

Одновременно развивается рынок специализированных услуг: проектирование киберустойчивых ИИ-архитектур и систем; тестирование безопасности ИИ; интеграция ИИ-инфраструктуры в классические центры мониторинга ИБ; аудит ИИ-систем; оценка и управление ИИ-рисками.

Весьма вероятно, что в ближайшие годы кибербезопасность ИИ станет таким же обязательным и базовым элементом корпоративной инфраструктуры цифровых компаний, каким сегодня являются защита сетей или облачных сред.

Источник: Gartner

НА ГЛОБАЛЬНОМ РЫНКЕ АКТИВНО РАЗВИВАЕТСЯ ЭКОСИСТЕМА СТАРТАПОВ И НОВЫХ РАЗРАБОТЧИКОВ В ОБЛАСТИ КИБЕРБЕЗОПАСНОСТИ ИИ-СИСТЕМ

Классы ИБ-решений и услуги в области безопасности ИИ	Описание	Примеры поставщиков решений и услуг
Защита ИИ-приложений (межсетевые экраны для ИИ)	Средства защиты ИИ-приложений и систем от манипуляции запросами, обхода ограничений, утечек данных и небезопасных ответов моделей	Lakera[1], Prompt Security[1], Lasso Security, PointGuard AI, 360 Digital Security Group (Китай)
Шлюзы безопасного доступа к ИИ и контроль использования ИИ	Средства централизованного контроля доступа к ИИ-сервисам, мониторинга использования ИИ, предотвращения несанкционированного применения ИИ сотрудниками и утечки данных	Prompt Security[1], Pangea[1], LayerX, Grip Security, PointGuard AI, Acronis (Сингапур)
Управление безопасностью моделей, ИИ-артефактов и цепочки поставок ИИ	Сканирование и контроль целостности ИИ-моделей, обнаружение уязвимостей в весах нейросетей, защита MLOps-конвейеров и репозиториев разработки от отравления данных и недокументированных возможностей (бэкдоров)	Protect AI, HiddenLayer, Robust Intelligence[1], PointGuard AI
Управление состоянием защищенности ИИ-систем	Средства инвентаризации, оценки конфигураций, выявления ошибок настройки, анализа поверхности атаки, контроля технических политик безопасности ИИ	Protect AI, Apex Security, Lakera[1], Noma Security, HiddenLayer, Cyata[1]
Безопасность ИИ-агентов, ИИ-систем в процессе выполнения	Контроль действий ИИ-агентов, доступа к инструментам, API, автономного поведения ИИ. Мониторинг, защита ИИ-систем в ходе выполнения, контроль запросов и аномалий	Noma Security, Prompt Security[1], Pangea[1], Apex Security, Lakera[1], PointGuard AI, Qi-Anxin (Китай)
Защита ИИ-систем, дополненных поиском, и ИИ-данных	Средства контроля доступа к базам знаний и механизмам поиска и извлечения данных, предотвращения утечек данных и компрометации ИИ-контекста	Lakera[1], PointGuard AI, Prompt Security[1], Pangea, Securiti
Наблюдаемость и мониторинг ИИ-систем	Средства журналирования, трассировки, мониторинга поведения ИИ-систем и расследования ИИ-инцидентов	Arize AI, Langfuse, WhyLabs, Fiddler AI, Baidu AI Cloud Observability (Китай)
Идентификация, управление правами доступа, политиками безопасности для ИИ	Средства управления учетными записями, полномочиями, политиками доступа и доверенными взаимодействиями ИИ-агентов и ИИ-сервисов	Astrix Security[1], Veza, StrongDM, Aembit, Huawei Cloud Security (Китай)
Управление ИИ-рисками	Средства оценки и управления ИИ-рисками, ведения реестра ИИ-моделей, - приложений и ИИ-систем, обеспечения соответствия требованиям регулирования, применения политик и контроля использования ИИ,	Credo AI, Holistic AI, Monitaur, PointGuard AI
Автоматизированное тестирование моделей и ИИ-приложений	Средства автоматизированного тестирования ИИ-приложений и систем на обход ограничений, утечки данных, небезопасного поведения и устойчивости к атакам	Mindgard, SPLXAI, Lakera Gandalf[1], HiddenLayer, PointGuard AI, 360 Digital Security Group (Китай)
Идентификация сгенерированного ИИ-контента	Средства обнаружения сгенерированных ИИ изображений, видео, голоса и документов, а также проверки происхождения цифрового контента	Reality Defender, Truepic, Pindrop, Hive, Baidu AI Content Security (Китай)
ИБ-услуги в области безопасности ИИ-систем	Проектирование киберустойчивых ИИ-систем, тестирование безопасности, интеграция ИИ-инфраструктуры в центры мониторинга ИБ, аудит ИИ-систем, управление ИИ-рисками	NCC Group, Trail of Bits, HiddenLayer, PointGuard AI, Mindgard, Protect AI, Akto, SecureIQLab

Источник: анализ исследовательской группы, список участников рынка неисчерпывающий и дан иллюстративно

1 – приобретены крупными глобальными поставщиками ИБ-решений

КРУПНЕЙШИЕ ГЛОБАЛЬНЫЕ РАЗРАБОТЧИКИ В СФЕРЕ ИБ ТАКЖЕ ВЫВОДЯТ НА РЫНОК РЕШЕНИЯ ДЛЯ ЗАЩИТЫ ИИ И ФОРМИРУЮТ ПЛАТФОРМЫ

Классы ИБ-решений и услуги в области безопасности ИИ	Примеры поставщиков решений и услуг
Защита ИИ-приложений (межсетевые экраны для ИИ)	Palo Alto, Check Point, Fortinet, Trend Micro
Шлюзы безопасного доступа к ИИ и контроль использования ИИ	Palo Alto, Microsoft, Zscaler, Netskope
Управление безопасностью моделей, ИИ-артефактов и цепочки поставок ИИ	Palo Alto, Trend Micro, Microsoft
Управление состоянием защищенности ИИ-систем	Palo Alto, Microsoft, Wiz (Google Cloud)
Безопасность ИИ-агентов, ИИ-систем в процессе выполнения	Palo Alto, Crowdstrike, Cisco, SentinelOne
Защита ИИ-систем, дополненных поиском, и ИИ-данных	Netskope, Microsoft
Наблюдаемость и мониторинг ИИ-систем	Crowdstrike, Cisco, SentinelOne
Идентификация, управление правами доступа, политиками безопасности для ИИ	CyberArk / Idira by PaloAlto, Okta, Cisco
Управление ИИ-рисками	Microsoft, IBM
Автоматизированное тестирование моделей и ИИ-приложений	Cisco
Идентификация сгенерированного ИИ-контента	Microsoft, Cisco, Check Point, Trend Micro
ИБ-услуги в области безопасности ИИ-систем	Accenture, EY, Deloitte, PwC

КОММЕНТАРИИ

Начиная с 2023 года крупнейшие игроки глобального рынка ИБ активно формируют собственные платформы безопасности ИИ, в среднем запустив от 3 до 5 и более продуктов в этой области.

Вендоры интегрируют защиту ИИ в существующие продукты и платформы (в том числе облачные, как AWS, Google Cloud), а также скупают специализированные стартапы:

- Palo Alto Networks – приобретение Protect AI
- SentinelOne – приобретение Prompt Security
- CrowdStrike – приобретение Pangea
- Check Point – приобретение Lakera, Cyata
- Cisco – приобретение Robust Intelligence, Astrix Security

Рынок быстро эволюционирует от защиты отдельных ИИ-моделей и приложений на базе больших языковых моделей к комплексным платформам защиты ИИ-систем во время выполнения, ИИ-агентов и автономной инфраструктуры ИИ.

Источник: анализ исследовательской группы, список участников рынка, сделок – неисчерпывающие и даны иллюстративно

РАЗВИТИЕ РЫНКА РЕШЕНИЙ КИБЕРЗАЩИТЫ ИИ МОЖЕТ ПОЙТИ 4 РАЗНЫМИ ПУТЯМИ, ВАЖНЫ ТЕНДЕНЦИИ НА ПЛАТФОРМЕННЫЕ РЕШЕНИЯ И ИНТЕРЕС К ОТКРЫТЫМ ИНСТРУМЕНТАМ ЗАЩИТЫ

« Рынок решений в сфере кибербезопасности ИИ пока зарождается, но уже можно отметить скорость его развития – в текущем году размер рынка превысит отметку 1% от всего глобального рынка ИБ. С высокой степенью уверенности можно также говорить, что рынок проходит схожую трансформацию, что ранее прошли рынки ИБ облачных технологий или мониторинга и реагирования на угрозы ИБ (SecOps): от набора точечных средств защиты к формированию полноценного платформенного слоя корпоративной безопасности.

Такой подход интегрирует защиту моделей, ИИ-агентов, данных и процессов исполнения ИИ в единый контур. Более того, по мере роста автономности ИИ-систем, рынок решений будет мигрировать от защиты ИИ-приложений к защите ИИ-агентов и далее к сквозной защите автономной ИИ-инфраструктуры и комплексному управлению рисками.

Одновременно в мире растет востребованность специализированных услуг по аудиту, тестированию и управлению безопасностью ИИ, а также построению киберустойчивых архитектур ИИ-систем.

Рынок только формируется, поэтому конечные победители и его будущая структура неясны. Пока можно сформулировать четыре возможных сценария или параллельные траектории развития:

1. Рынок сконцентрируется вокруг крупных разработчиков ИБ-решений, как это произошло с облачной ИБ, EDR/XDR.
2. Появится отдельный лидер или лидеры рынка.
3. Безопасность ИИ постепенно станет частью безопасности приложений, облаков, защиты данных, управления идентификацией и доступом и превратится из отдельного рынка во встроенную функцию существующих платформ кибербезопасности.
4. Поставщики ИИ-систем и платформ ИИ-разработки сами формируют или приобретают продукты встроенной кибербезопасности, подобно тому, как Microsoft, VMware и другие вендоры встраивают ИБ-решения в пакеты своих программных продуктов.

Кроме того, на глобальном рынке наблюдается устойчивый рост интереса к открытым инструментам и решениям защиты систем ИИ с открытым исходным кодом: инструменты тестирования безопасности моделей (Garak, PyRIT, AgentDojo), системы защиты и контроля больших языковых моделей (NVIDIA NeMo Guardrails), средства мониторинга и наблюдаемости ИИ-систем (Langfuse, Arize Phoenix), а также подходы к моделированию угроз и базы знаний об атаках на ИИ (MITRE ATLAS) »

Сергей Салов
Партнер Группы компаний Б1, стратегическое консультирование компаний сектора телеком, медиа и высоких технологий

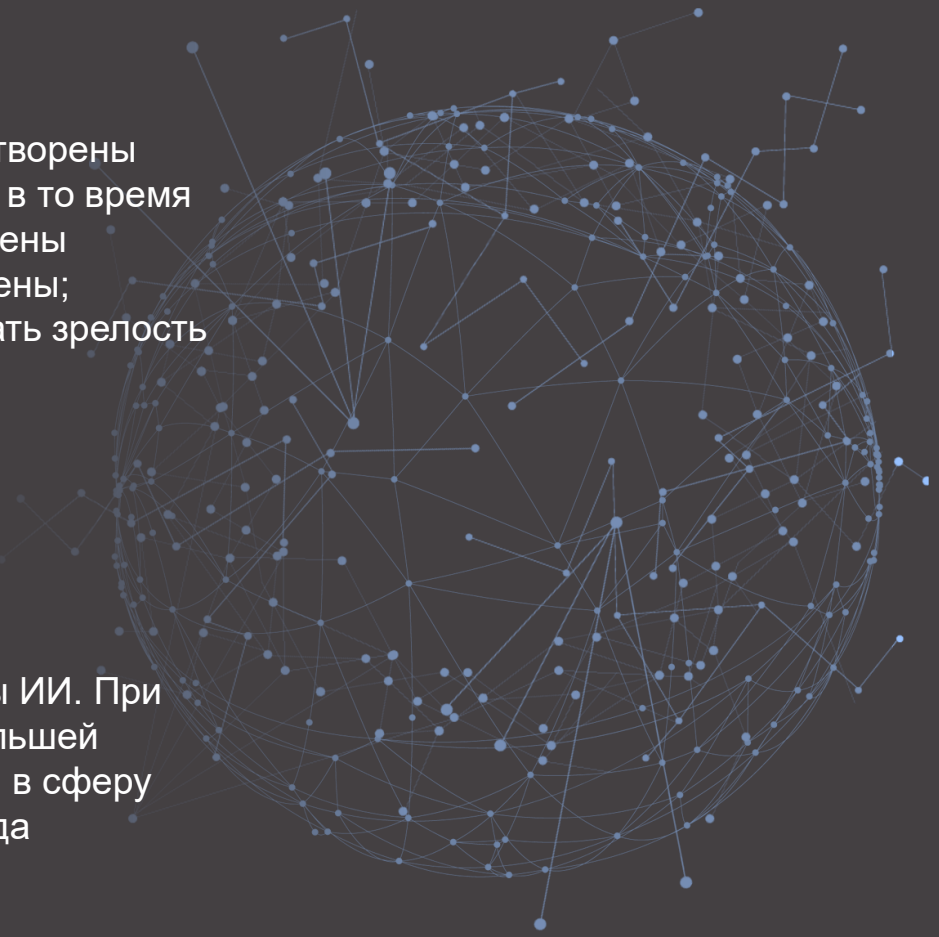
ЧТОБЫ ПРОАНАЛИЗИРОВАТЬ СОСТОЯНИЕ РЫНКА УСЛУГ И СРЕДСТВ ЗАЩИТЫ ИИ В РОССИИ, МЫ ОПРОСИЛИ ПОРЯДКА 60 ОТЕЧЕСТВЕННЫХ КОМПАНИЙ

~60 респондентов

из числа крупных компаний разных отраслей стали участниками опроса; больше всего представителей финансовой сферы и страхования, промышленности, ритейла и электронной коммерции, ИТ

41% компаний

полностью или скорее удовлетворены текущим уровнем защиты ИИ, в то время как 40% скорее не удовлетворены или полностью не удовлетворены; еще 19% считают, что оценивать зрелость защиты пока рано



23% компаний

отмечают ИИ среди топ-3 стратегических приоритетов, **12%** называют ИИ критически важным для ключевых процессов; еще **46%** считают ИИ значимым для отдельных процессов

49% компаний

выделяют бюджет для защиты ИИ. При этом **53%** ожидают рост (в большей части умеренный) инвестиций в сферу защиты ИИ в ближайшие 2 года

КАК И ПОЧЕМУ ВОЗНИКЛА ИДЕЯ ЭТОГО ИССЛЕДОВАНИЯ

« Сейчас рынок внедрения ИИ находится в необычной точке. Многие компании уже приняли решение искать эффективность в новых ИИ-системах и запускать пилотные проекты. Мы наблюдаем, как в бизнесе появились первые чат-боты и ассистенты, а сотрудники многих компаний получили доступ к широкому набору нейросетей для рабочих задач.

Как и в случае с другими технологиями, внедрение ИИ неизбежно связано с рисками. Сегодня это прежде всего утечки данных, теневого и неконтролируемого использования ИИ и вопрос контролируемости ответов моделей в соответствии с политиками компании. Однако уже сейчас мы видим переход к следующему этапу – к более автономным системам.

Хороший пример – кодинговые агенты, которые не просто помогают писать код, а становятся важной частью процесса разработки и уже могут самостоятельно выполнять отдельные задачи.

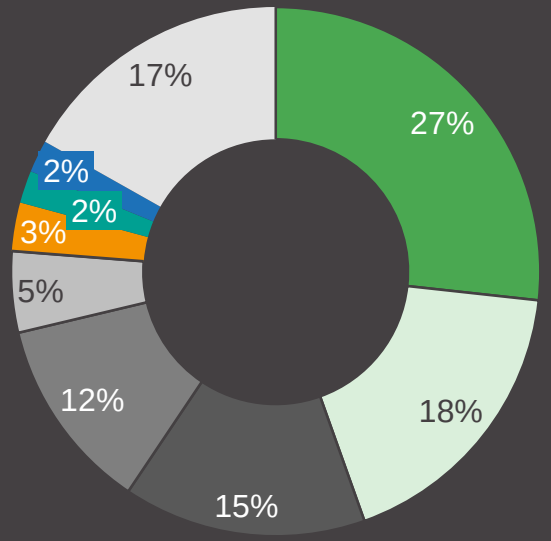
По мере того как такие агентные системы будут проникать в разные бизнес-функции компаний и использоваться в ежедневных процессах, критически важными станут их контроль и защита. Нам нужно будет понимать, какие действия они выполняют, насколько устойчиво работают, а также какие риски они создают для данных, процессов и инфраструктуры.

Главная ценность настоящего исследования, на мой взгляд, заключается в том, что оно раскрывает, как ключевые участники российского рынка видят безопасность новых технологий на текущем этапе. Отечественные компании адаптируются к ИИ в непростой среде, с ограничениями по вычислительным мощностям и доступности моделей. Поэтому особенно важно понять, какие риски бизнес уже считает значимыми и какой уровень зрелости формируется на рынке ».

Евгений Кокуйкин
Основатель HiveTrace

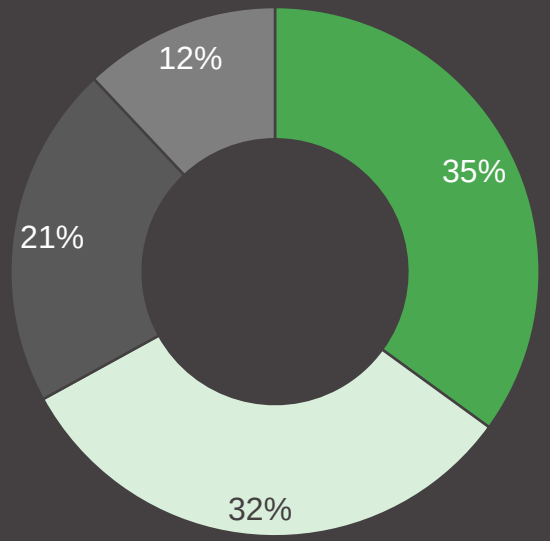
В ОПРОСЕ ПРИНИМАЛИ УЧАСТИЕ РЕСПОНДЕНТЫ ИЗ БОЛЕЕ ЧЕМ 10 ОТРАСЛЕЙ, 2/3 КОМПАНИЙ С ЧИСЛЕННОСТЬЮ БОЛЕЕ 2000 СОТРУДНИКОВ

Отраслевая структура респондентов,
% от общих ответов



- Финансы / банки / финтех и страхование
- Промышленность / производство
- Ритейл / элетронная коммерция
- ИТ
- Транспорт / логистика
- Телеком / медиа
- Образовние / наука
- Гос сектор
- Прочие отрасли

Структура респондентов по размеру компаний,
% от общих ответов



- 10 000 и более сотрудников
- 2 000 – 9 999 сотрудников
- Политик контроля нет, или они не доведены до сотрудников
- Менее 500 сотрудников

КОММЕНТАРИИ

- В опросе приняли участие **≈60** респондентов из разных отраслей.
- Больше всего респондентов из таких отраслей, как финансовый сектор, промышленность, ритейл и электронная коммерция.
- Основу панели составляют профильные ИБ-роли: **39%** – CISO, **16%** – руководители направлений ИБ, **12%** – архитекторы / инженеры.
- В панели представлены роли, связанные с ИИ и данными: **14%** – руководители направлений ИИ / данных, что позволяет учитывать потребности лидеров ИИ-инициатив.
- Большинство респондентов (**67%**) представляют крупные компании с численностью от **2000** сотрудиков, что повышает релевантность выводов для корпоративного рынка.

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

ИИ АКТИВНО ВНЕДРЯЕТСЯ, КОМПАНИИ УЖЕ СТАЛКИВАЮТСЯ С ИБ-ИНЦИДЕНТАМИ, 40% РЕСПОНДЕНТОВ НЕ УДОВЛЕТВОРЕНЫ ЗАЩИТОЙ ИИ-СИСТЕМ

- Результаты опроса показывают, что крупные отечественные компании активно инвестируют в технологии ИИ: **97%** респондентов используют, пилотируют ИИ или считают его значимым для бизнеса.
- **35%** опрошенных считают ИИ топ-3 стратегическим приоритетом или критически важной технологией для ключевых процессов, еще **46%** выделяют ИИ как важный элемент для отдельных бизнес-процессов.
- Наиболее распространенные типы используемого ИИ – большие языковые модели / ГенИИ, ИИ-системы дополненные поиском и ИИ- агенты. ИИ на базе больших языковых моделей используют **79%** опрошенных компаний, ИИ-системы, дополненные поиском (RAG), – **67%**, почти половина опрошенных компаний использует ИИ-агентов с доступом к системам – **46%**.
- При этом уровень удовлетворенности защитой ИИ остается неоднозначным: **41%** компаний полностью или скорее удовлетворены текущим уровнем защиты. Еще **33%** скорее не удовлетворены, а **7%** полностью не удовлетворены уровнем защиты; **19%** считают, что оценивать зрелость пока рано.
- Организации уже сталкиваются с инцидентами, связанными с использованием ИИ: такие случаи отметили **19%** респондентов.

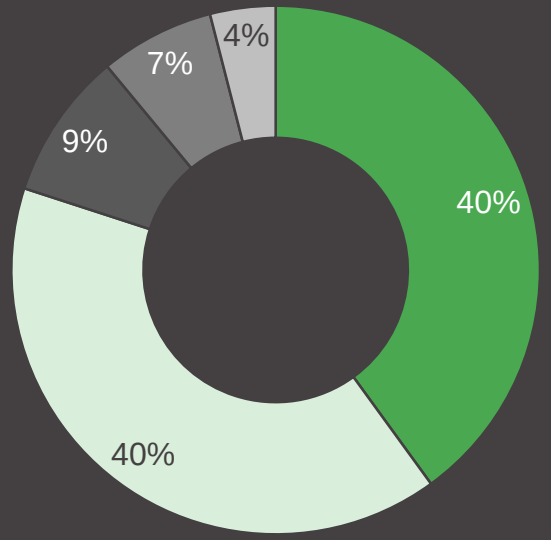
Топ-10 направлений использования ИИ,
% от общих ответов

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

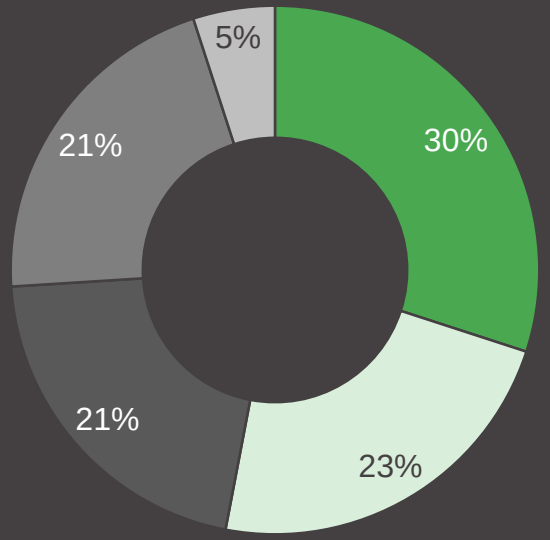
ОРГАНИЗАЦИИ АКТИВНО РАЗВОРАЧИВАЮТ ИИ В ЧАСТНОЙ ИНФРАСТРУКТУРЕ, 42% НЕ ОБЕСПЕЧИВАЮТ КОНТРОЛЬ ИСПОЛЬЗОВАНИЯ ИИ СОТРУДНИКАМИ

Основные модели развертывания ИИ, % от общих ответов



- В закрытом / частном контуре
- Гибридный подход
- Во внешних ИИ-сервисах с ограничениями
- Во внешних ИИ-сервисах без ограничений
- Подход пока не определен

Контроль «теневого использования» ИИ сотрудниками, % от общих ответов



- Разрешено с мониторингом и контролем
- Разрешены только утвержденные ИИ- сервисы
- Политик контроля нет, или они не доведены до сотрудников
- Формально разрешено, без технического контроля
- Полностью запрещено и технически блокируется

КОММЕНТАРИИ

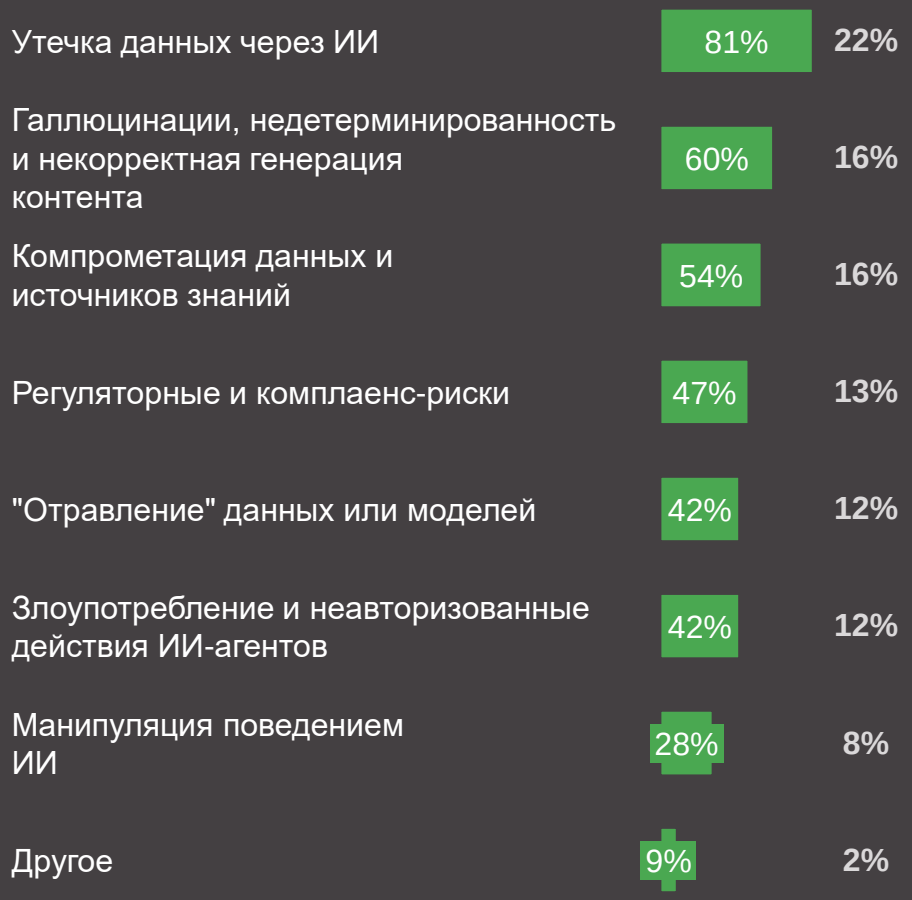
- **80%** респондентов используют закрытый или гибридный контур, что указывает на стремление ограничивать передачу данных во внешние ИИ-сервисы и системы.
- **42%** компаний находятся в слабой зоне контроля использования ИИ сотрудниками: политики отсутствуют, не доведены до сотрудников или использование ИИ разрешено без технического контроля.
- Контроль теневого использования входит в приоритеты на ближайший год для **33%** респондентов. Это подтверждает, что неконтролируемое использование ИИ становится отдельным направлением защиты.
- Функцию контроля доступа и блокировки сторонних неавторизованных ИИ-инструментов считают ценной **30%** респондентов.

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

УЧАСТНИКИ ОПРОСА ПРИЗНАЮТ КЛЮЧЕВЫЕ КИБЕРРИСКИ ИИ, УЖЕ НАЧИНАЮТ РЕАЛИЗОВЫВАТЬ КОМПЛЕКС МЕР ПО ЗАЩИТЕ ИИ, НО НЕ В ПОЛНОМ МАСШТАБЕ

Наиболее критичные риски для ИИ-систем,
% от респондентов и % от общего числа ответов



Топ-10 мер защиты ИИ, которые уже реализуются в организациях,
% от респондентов

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

ОСНОВНЫЕ ПРИОРИТЕТЫ ЗАЩИТЫ ИИ НА БЛИЖАЙШУЮ ПЕРСПЕКТИВУ

- **60%** компаний измеряют эффективность мер по защите ИИ
- **46%** респондентов используют ИИ-агентов с доступом к системам, при этом **37%** респондентов осуществляют контроль работы автономных ИИ-решений и ИИ-агентов (действия, инструменты, API)
- **25%** респондентов имеют политики ИБ для ИИ, **44%** считают это приоритетным направлением
- **Среди основных приоритетов защиты ИИ на ближайший год (% от числа опрошенных компаний):**

1. Использование ИИ в закрытом контуре (**46%**)
2. Защита данных и знаний (**46%**)
3. Формирование стратегии и политик защиты ИИ (**44%**)
4. Контроль «теневого использования» ИИ сотрудниками (**33%**)
5. Защита ИИ-агентов (**26%**)
6. Мониторинг и реагирование / SOC (**26%**)

ВЗГЛЯД ЭКСПЕРТА И УЧАСТНИКА РЫНКА

« Мы строим эшелонированную защиту вокруг использования ГенИИ. Базовый уровень – защита ГенИИ приложений – модульное решение, которое маскирует конфиденциальные данные, блокирует нежелательную передачу вовне и противодействует промт-инъекциям. Понимаем, что точечной фильтрации недостаточно, поэтому активно развиваем мониторинг и поведенческий скоринг сотрудников. Это даст прозрачную картину и позволит выявлять систематические попытки обхода правил.

Следующий вызов, к которому мы готовимся уже сейчас, – рост автономности ИИ-агентов. Если раньше речь шла в основном о защите информации, то завтра агенты смогут выполнять реальные действия: от удаления данных до изменений в инфраструктуре. В индустрии уже фиксируются инциденты, когда ассистенты некорректно интерпретировали команды и наносили ущерб. Это формирует новый класс угроз, и мы фокусируемся на проактивных мерах – закладываем защиту от подобных сценариев на уровне архитектуры наших решений »

Денис Султанов

Руководитель управления безопасности приложений, БКС

Источник: результаты анкетирования

КИБЕРРИСКИ ДЛЯ ИИ И ПРИМЕНЯЕМЫЕ МЕРЫ ЗАЩИТЫ: ВЫВОДЫ

ИИ уже воспринимается как значимый элемент цифровой трансформации бизнеса. Для **23%** респондентов ИИ входит в число трех главных стратегических приоритетов организации, еще для **12%** он критически важен в ключевых бизнес-процессах. Это означает, что более трети организаций уже находятся в зоне, где нарушения безопасности ИИ способны оказывать прямое влияние на бизнес-результаты.

Наиболее распространенные сценарии применения – это внутренние корпоративные помощники, поиск по базам знаний, автоматизация обработки документов, клиентские чат-боты, аналитика и поддержка разработки программного обеспечения. Среди используемых технологий доминируют большие языковые модели, ИИ-системы с дополнительным поиском по корпоративным базам (RAG) и классические модели машинного обучения. При этом уже **46%** опрошенных организаций используют ИИ-агентов с доступом к корпоративным системам, что существенно расширяет потенциальную поверхность атак.

В архитектурном плане рынок демонстрирует осторожный подход: значительное число систем размещается в частном контуре.

Наиболее значимыми рисками для собственных ИИ-систем респонденты считают утечку данных через ИИ, галлюцинации и некорректную генерацию контента, компрометацию источников знаний и данных, регуляторные риски, «отравление» данных и моделей, а также несанкционированные действия ИИ-агентов.

Только **30%** организаций разрешают использование ИИ с мониторингом и контролем, тогда как в **42%** компаний либо отсутствует технический контроль, либо отсутствуют политики использования ИИ вообще. Это указывает на разрыв между фактическим использованием ИИ сотрудниками и существующими механизмами управления рисками.

Исследование также выявило разрыв между уровнем внедрения технологий ИИ и уровнем внедрения специализированных мер их защиты. Так, ИИ-системы с поиском по корпоративным базам используют две трети респондентов, однако меры по защите источников знаний внедрены лишь у **30%**.
ИИ-агентов применяют **46%** респондентов, тогда как специализированный контроль их действий реализован только у **37%**.

Большие языковые модели применяют более чем ¾ респондентов, в то же время механизмы защиты данных внедряют только порядка **50%**, шлюзы, фильтрацию и контроль запросов для ИИ – менее половины, мониторинг поведения моделей и тестирования их устойчивости реализованы лишь у примерно четверти организаций.

Несмотря на высокий уровень осознания рисков, практическая зрелость киберзащиты ИИ систем остается ограниченной. Только у **25%** организаций уже существует формализованная политика безопасности ИИ, тогда как у большинства (**53%**) она находится на стадии разработки. Наиболее распространенные меры защиты связаны с контролем использования ИИ сотрудниками, управлением доступом, обучением персонала и защитой данных.

В целом результаты исследования показывают, что российские организации уже рассматривают ИИ как важный бизнес-инструмент и осознают связанные с ним риски. Однако внедрение технологий ИИ опережает развитие механизмов их защиты. Наиболее выраженные дефициты наблюдаются в областях **защиты больших языковых моделей, защиты систем с поиском по корпоративным данным и безопасности ИИ-агентов**. Именно эти направления, вероятно, станут основными драйверами развития рынка услуг и решений кибербезопасности ИИ в ближайшие годы.

Источники: анализ исследовательской группы, результаты анкетирования

МНЕНИЕ ЭКСПЕРТОВ И УЧАСТНИКОВ РЫНКА

« Результаты исследования довольно хорошо показывают реальную картину: ИИ уже фактически используется в бизнес-процессах, но безопасность за этим пока не успевает. Почти половина компаний считает ИИ важным для отдельных процессов, а для части он уже становится критичным или стратегическим направлением. При этом полностью довольны текущим уровнем защиты ИИ всего 9%, а 40% прямо говорит, что скорее и полностью недовольны или находятся в поиске решения.

Для меня здесь главный вывод такой: ИИ уже уходит в продакшн, но во многих компаниях это происходит быстрее, чем появляется стратегия его защиты. То есть бизнес уже использует модели, сервисы, ассистентов, автоматизацию, а вопросы вроде ответственности, контроля, тестирования, мониторинга и оценки рисков часто остаются на втором плане.

Хорошо, что рынок уже видит проблему. Например, более 40% компаний рассматривают внедрение специализированных отечественных решений для защиты ИИ. Но пока это часто выглядит как отдельные инициативы: где-то что-то проверили, где-то купили инструмент, где-то начали обсуждать редтиминг. А системного подхода во многих случаях еще нет.

На мой взгляд, разрыв можно сократить, только если смотреть на безопасность ИИ не как на один инструмент, а как на полноценную управленческую и техническую задачу. Нужны стратегия, понятные зоны ответственности, обучение команд, инвентаризация ИИ-сценариев и регулярная проверка ИИ-систем на устойчивость. Иначе безопасность все время будет идти следом за внедрением, а в случае с ИИ это слишком медленная позиция »

Александр Шепилов

Директор направления обеспечения информационной безопасности, Ростелеком

« Мы видим, что основной запрос бизнеса в сфере кибербезопасности ИИ пока сфокусирован вокруг трех направлений: **защита данных от последствий использования ИИ**; **защита самого ИИ от нарушителя** (отравление данных и промпт-инъекции); **защита от атакующего ИИ**.

В части защиты от последствий использования ИИ мы сформулировали для себя 5 базовых правил 1. Идентичность: ИИ-агент должен иметь четкую роль, права, ограниченный круг доступа. 2. Классификация данных и DLP. 3. Отслеживаемость: логирование всех действий и решений модели 4. Негативные тесты: проверка ИИ перед допуском в контур на промпт-инъекции, попытки утечки через контекст и галлюцинации. 5. Архитектурный паттерн: проектирование систем так, чтобы исключить риски компрометации данных.

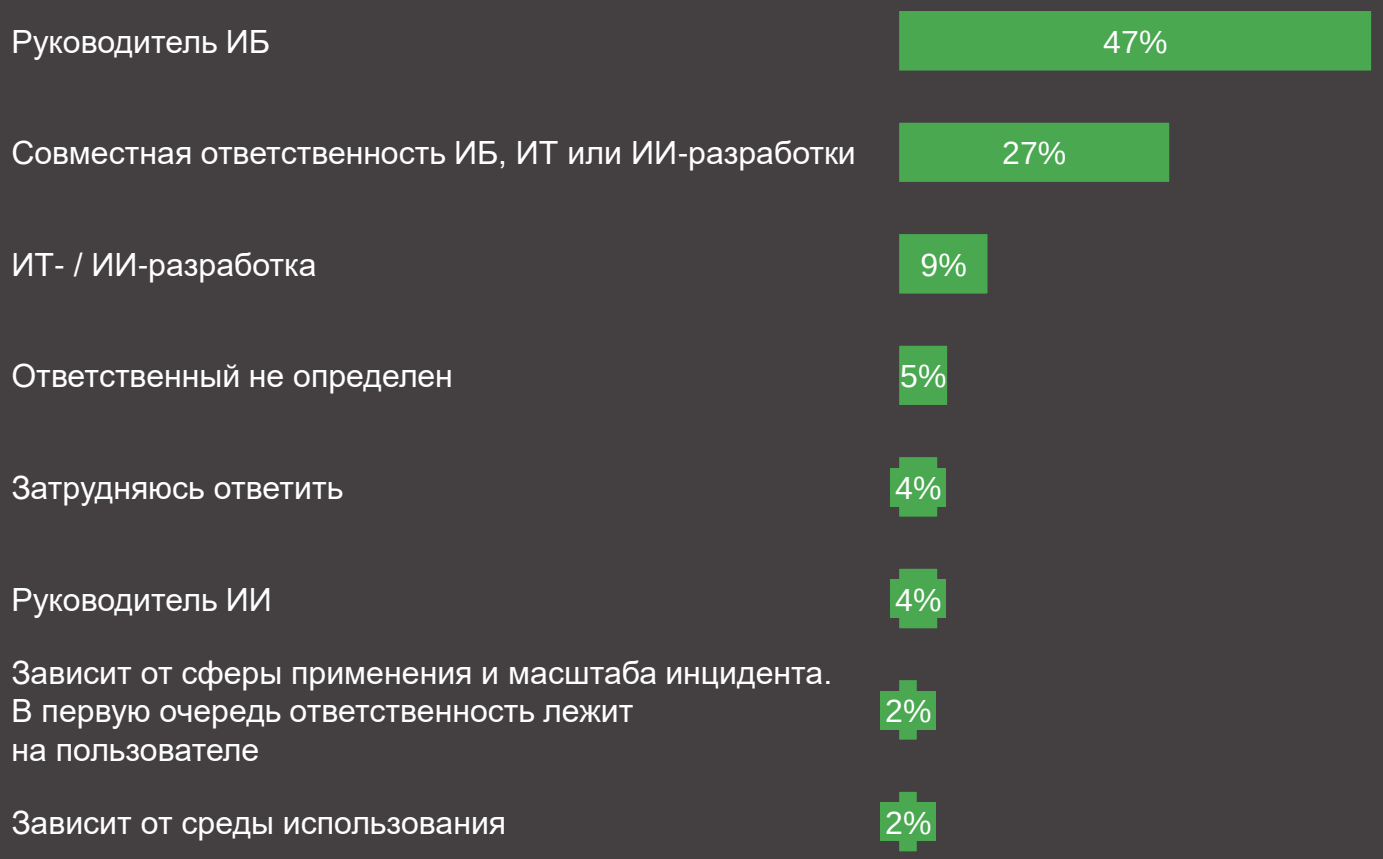
Кроме того, бизнесу нужно научиться по-новому решать несколько классов задач. Это безопасная трансформация процессов разработки ПО: («вайбкодинг» и агентсткая разработка), внедрение правил общения с публичными нейросетями, а также контроль трафика с помощью веб-шлюзов безопасности. Отдельно стоит задача защиты ИИ-моделей от «отравления» данных. Наконец, **контроль ИИ-агентов внутри компании** – это новая сущность. Появляется новый субъект, который наравне с пользователями взаимодействует с бизнес-системами. Важно разграничивать их права и отслеживать, какие учетные записи агент использует, куда имеет доступ, какой информацией обменивается и какие активности выполняет внутри сессии »

Иван Вассунов

Директор Центра технологий кибербезопасности, ГК «Солар»

КОМПАНИИ В РФ ПРИНИМАЮТ ОРГАНИЗАЦИОННЫЕ МЕРЫ ДЛЯ ЗАЩИТЫ ИИ: НАЗНАЧАЮТ ОТВЕТСТВЕННЫХ, ВЫДЕЛЯЮТ БЮДЖЕТ, НАРАЩИВАЮТ ИНВЕСТИЦИИ

Ответственный за киберзащиту систем ИИ в организации,
% от общих ответов



- **49%** компаний выделяют бюджет на защиту ИИ[1]. Наиболее высока доля компаний с бюджетами на защиту ИИ среди организаций, которые считают ИИ одним из трех главных стратегических приоритетов и критически важным для ключевых процессов. Более **40%** опрошенных не выделяют соответствующий бюджет или затруднились с ответом.
- При этом **15%** опрошенных компаний, где ИИ считается стратегическим приоритетом, не выделяют бюджеты на защиту ИИ.
- Среди компаний, для которых ИИ важен в отдельных процессах, бюджет на защиту ИИ выделяется лишь в **42%** случаев.
- **53%** опрошенных планируют увеличить траты на защиту ИИ, **около четверти** (**23%**) не планируют инвестиции или хотят оставить их на текущем уровне; лишь **5%** планируют их сократить.
- Наиболее популярная модель ответственности – **47%** возлагают защиту систем ИИ на руководителя ИБ. В ряде случаев ответственность может быть размыта: в **9%** случаев ответственный не определен или его затруднительно обозначить; у **4%** респондентов ответственность зависит от контекста.

Источник: результаты анкетирования 1 – отдельный бюджет или в составе ИТ- или ИБ-бюджетов

РОССИЙСКИЙ РЫНОК РЕШЕНИЙ И УСЛУГ КИБЕРБЕЗОПАСНОСТИ ИИ НАХОДИТСЯ НА СТАРТОВОЙ ТРАЕКТОРИИ РАЗВИТИЯ

Российский рынок киберзащиты систем искусственного интеллекта находится на начальной стадии развития.

Несмотря на активное использование больших языковых моделей, систем с поиском по корпоративным данным и интеллектуальных агентов, специализированные средства защиты ИИ пока ограниченно распространены. Согласно результатам исследования, только **16%** организаций уже используют решения внешних поставщиков для защиты ИИ-систем.

При этом основной моделью обеспечения безопасности ИИ остаются внутренние разработки: **52%** компаний используют собственные механизмы защиты и компетенции. Такая ситуация объясняется как относительной незрелостью рынка, так и ограниченным предложением специализированных отечественных продуктов. Одновременно еще 41% организаций планируют внедрение внешних решений, что свидетельствует о постепенном формировании спроса на данные классы решений, особенно при дальнейшем масштабировании использования ИИ.

В последние два года в России начал формироваться класс специализированных поставщиков*: HiveTrace, InfEra Security, Strazh, AppSec Solutions и др. На рынок выходят платформы корпоративных разработчиков, например SOWA AI от СберТех.

Параллельно формируется рынок услуг по безопасности ИИ. Консалтинговые и сервисные направления развивают такие компании*, как Swordfish Security, BI.ZONE, ГК «Солар», Infosecurity (Softline), Kaspersky (Kaspersky AI Security Team) и другие игроки рынка. Их предложения включают аудит защищенности ИИ-систем, оценку рисков, тестирование безопасности, разработку политик безопасного использования ИИ, сопровождение внедрения технологий ИИ в корпоративной среде.

В целом рынок решений и услуг в сфере кибербезопасности ИИ в России находится на этапе перехода от отдельных инициатив и пилотных проектов к формированию полноценного сегмента.

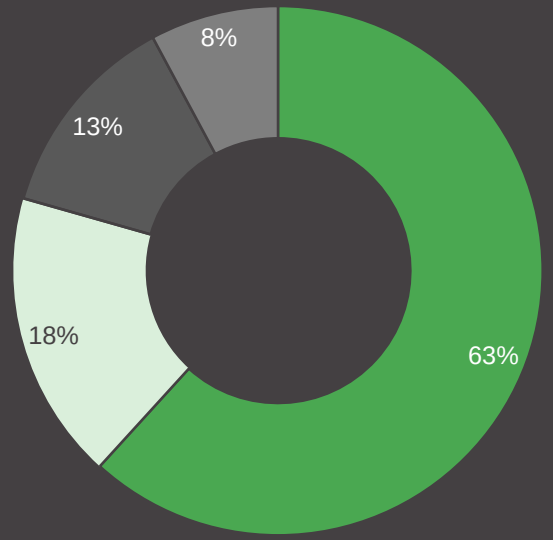
Источники: анализ исследовательской группы

* Список участников рынка неисчерпывающий и дан иллюстративно

ПРИМЕНЕНИЕ ТИРАЖНЫХ РЕШЕНИЙ ПО КИБЕРЗАЩИТЕ ИИ ПОКА ОГРАНИЧЕНО, ОСНОВНАЯ МОДЕЛЬ ЗАЩИТЫ – ВНУТРЕННИЕ РАЗРАБОТКИ И КОМПЕТЕНЦИИ

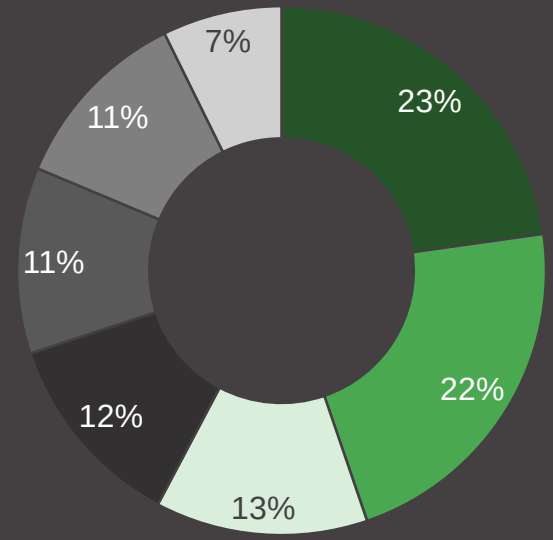
- Только **16%** респондентов уже применяют специализированные решения от внешних разработчиков для обеспечения безопасности ИИ.
- При этом **46%** готовы их использовать в ближайшем времени, хотя на данный момент не применяют.
- Уровень использования внутренних разработок для защиты ИИ составляет **51%** – сейчас это основные инструменты защиты.
- При этом **19%** респондентов не применяют сейчас и не рассматривают такие решения на ближайшее время.

Оценка зрелости доступных на рынке ИБ-решений для защиты ИИ, % от общих ответов



- Встречаются отдельные решения
- Не знаем о готовых специализированных решений для безопасности ИИ от внешних разработчиков
- На рынке отсутствуют необходимые решения
- Множество зрелых решений в большом спектре

Основные проблемы ИБ-решений для защиты ИИ, % от общих ответов



- Отсутствие практической апробации
- Недостаток специализированных отечественных решений
- Высокая стоимость
- Отсутствие сквозной защиты ИИ
- Низкий уровень маркетинговой активности
- Сложная интеграция и внедрение
- Другое

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

НАИБОЛЕЕ ВОСТРЕБОВАНЫ ФУНКЦИИ ЗАЩИТЫ ДАННЫХ, КОНТРОЛЬ ИСПОЛЬЗОВАНИЯ ИИ, СРЕДИ ТОП-5 ПРИОРИТЕТОВ – ЗАЩИТА ИИ-АГЕНТОВ

- Уровень зрелости рынка отечественных киберрешений для защиты ИИ пока оценивается как невысокий. Лишь **8%** респондентов считают, что на рынке представлено достаточно зрелых решений для защиты ИИ. **45%** отмечают наличие лишь отдельных продуктов, а значительная часть респондентов вообще не знакома со специализированными решениями данного класса.
- Наиболее востребованные функции – это предотвращение утечек данных через ИИ (**59%**), автоматическая защита персональных и конфиденциальных данных (**55%**), защита от атак через запросы к моделям (**41%**), контроль использования несанкционированных ИИ-сервисов (**30%**) и защита интеллектуальных агентов (**27%**). Значительно меньший интерес вызывают задачи по защите моделей от копирования и цифровой маркировки контента, что свидетельствует о концентрации внимания прежде всего на защите данных и снижении операционных рисков.
- В целом результаты исследования показывают, что рынок переходит от стадии осознания рисков к стадии практического внедрения средств защиты. При относительно низком текущем уровне использования специализированных решений (16%) потенциальный интерес к ним в три раза выше. Одновременно формируется запрос на отечественные легкоинтегрируемые и комплексные платформы, обеспечивающие защиту ИИ на протяжении всего жизненного цикла.

Топ-10 наиболее ценных функций во внешних решениях для безопасности ИИ,
% от количество респондентов

Источник: результаты анкетирования

КИБЕРБЕЗОПАСНОСТЬ ИИ СТАНОВИТСЯ ВАЖНЫМ КОМПОНЕНТОМ ДОВЕРЕННОГО И БЕЗОПАСНОГО ИСПОЛЬЗОВАНИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Настоящее исследование посвящено вопросам **киберзащиты систем искусственного интеллекта**. При этом важно различать понятия «**безопасность искусственного интеллекта»** и «**киберзащита систем искусственного интеллекта»**, поскольку они описывают разные, но взаимосвязанные категории рисков.

Под безопасностью искусственного интеллекта понимается обеспечение безопасного, надежного и контролируемого поведения ИИ-систем, предотвращение причинения вреда человеку, обществу или инфраструктуре вследствие ошибок модели, предвзятости, галлюцинаций, некорректного целеполагания или потери управляемости.

Под киберзащитой систем искусственного интеллекта понимается защита ИИ-систем от злонамеренного воздействия, включая атаки на модели, компрометацию данных, «отравление» обучающих выборок, состязательные атаки, атаки через инструкции, кражу моделей и компрометацию ИИ-агентов.

Несмотря на различие целей, защита ИИ – это неотъемлемая часть безопасности искусственного интеллекта.

Во многих случаях нарушение защищенности модели непосредственно приводит к нарушению ее безопасного функционирования. Например, «отравление» обучающих данных может изменить поведение модели и привести к ошибочным решениям, состязательные атаки способны вызвать некорректную классификацию объектов в критически важных системах, а атаки через инструкции позволяют обходить встроенные ограничения больших языковых моделей. По этой причине современные концепции **доверенного искусственного интеллекта** рассматривают защиту ИИ как один из обязательных элементов обеспечения его безопасности наряду с надежностью, прозрачностью, объяснимостью и управляемостью.

За последние несколько лет регулирование искусственного интеллекта во всем мире сместилось от общих принципов «ответственного ИИ» к практическим требованиям по управлению рисками, устойчивости и кибербезопасности ИИ-систем. Одновременно происходит сближение регулирования ИИ и традиционной информационной безопасности.

Китай развивает регулирование через специализированные нормативные акты и технические стандарты. Базовым документом являются «**Временные меры по управлению сервисами генеративного искусственного интеллекта»**, устанавливающие требования к оценке безопасности, защите данных и управлению рисками. Дополнительно развиваются национальные стандарты и рекомендации по управлению безопасностью ИИ в рамках деятельности технического комитета TC260.

В США ключевыми документами являются **Указ Президента США № 14110 «О безопасной, защищенной и заслуживающей доверия разработке и использовании искусственного интеллекта»** и разработанная Национальным институтом стандартов и технологий США «**Система управления рисками искусственного интеллекта»**. Эти документы закрепляют риск-ориентированный подход, необходимость тестирования безопасности моделей, оценки угроз и внедрения мер контроля на протяжении всего жизненного цикла ИИ-систем.

Источники: анализ исследовательской группы

ПОСТЕПЕННО РАЗВИВАЕТСЯ СИСТЕМА РЕГУЛИРОВАНИЯ, СПЕЦИАЛИЗИРОВАННЫХ СТАНДАРТОВ И МЕТОДИЧЕСКИХ ФРЕЙМВОРКОВ

В Европейском союзе основу регулирования составляет **Регламент Европейского союза об искусственном интеллекте**. Документ устанавливает обязательные требования для высокорисковых ИИ-систем и прямо включает кибербезопасность в перечень требований к их надежности и устойчивости. В европейской модели защита ИИ фактически рассматривается как составная часть обеспечения безопасности ИИ.

В России специализированное регулирование защиты ИИ находится на стадии формирования. В настоящее время требования к безопасности ИИ распространяются через действующую систему регулирования информационной безопасности. Ключевое значение имеет **Приказ ФСТЭК России от 11 апреля 2025 г. № 117**, устанавливающий современные требования к защите государственных информационных систем в части применения ИИ. Параллельно развивается система национальных стандартов по искусственному интеллекту, включая **ГОСТ Р 59276-2020 «Системы искусственного интеллекта. Способы обеспечения доверия»** и стандарты **Технического комитета 164 «Искусственный интеллект».**

Дальнейшим существенным этапом стало принятие Федерального закона **«О поддержке развития технологий искусственного интеллекта в РФ»**. Документ закрепил понятия в сфере ИИ, принципы госрегулирования, меры поддержки и особенности применения больших фундаментальных моделей. Закон также создал правовую основу для требований к безопасности, доверию и маркировке ИИ-контента. При этом он предполагает д дальнейшую детализацию через акты Правительства РФ и ведомств.

Параллельно с регулированием формируется международная система специализированных стандартов и методических фреймворков. Наибольшее значение в этом контексте имеют **Система управления рисками искусственного интеллекта** (NIST AI RMF), **Профиль управления рисками генеративного искусственного интеллекта** (NIST AI 600-1 Generative AI Profile), **База знаний по угрозам для систем искусственного интеллекта** (MITRE Adversarial Threat Landscape for Artificial-Intelligence Systems, MITRE ATLAS), **Десять наиболее критичных угроз для приложений на основе больших языковых моделей** (OWASP Top 10 for LLM Applications), а также международные стандарты **«Система менеджмента искусственного интеллекта»** (ISO/IEC 42001) и «**Управление рисками искусственного интеллекта»** (ISO/IEC 23894). Важную роль играет **MITRE ATLAS**, который становится отраслевой базой знаний по угрозам и атакам на ИИ-системы, аналогичной MITRE ATT&CK в традиционной кибербезопасности. В России также **формируются собственные фреймворки безопасности ИИ** с учетом особенностей национальных задач, политик и регулирования (AI-SAFE, SAIMM и др.)

Таким образом, мировая тенденция заключается в признании того, что защита систем искусственного интеллекта – самостоятельное направление кибербезопасности, требующее специализированных моделей угроз, методов тестирования и средств защиты.

Безопасное функционирование систем ИИ невозможно без его защищенности от преднамеренного воздействия. Именно поэтому современные регуляторные инициативы, стандарты и отраслевые фреймворки все чаще рассматривают безопасность и защищенность искусственного интеллекта как единую систему требований к доверенному и безопасному созданию и применению искусственного интеллекта.

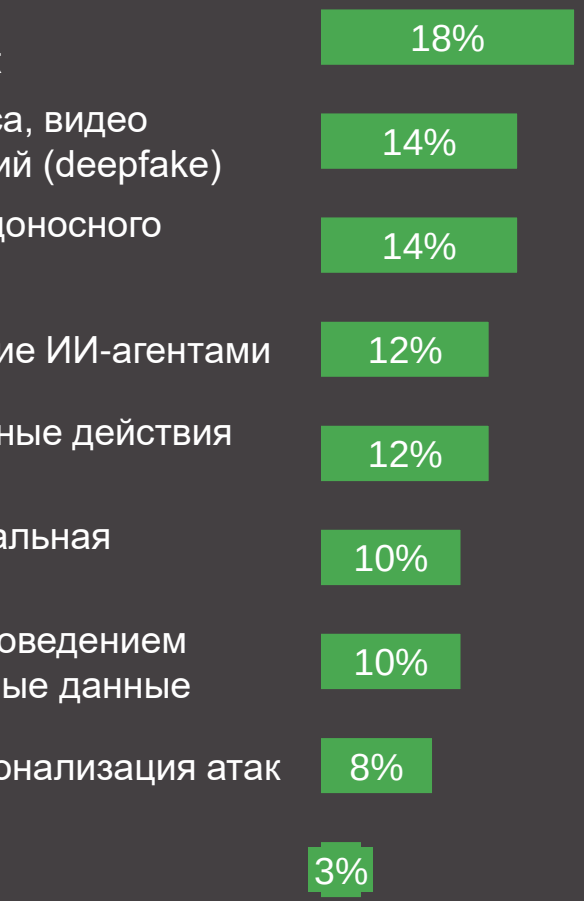
Источники: анализ исследовательской группы

БИЗНЕС КАК В РОССИИ, ТАК И ГЛОБАЛЬНО УЖЕ ОСОЗНАЕТ, ЧТО ИИ МОЖЕТ СТАТЬ ЗНАЧИМЫМ ИСТОЧНИКОМ КИБЕРУГРОЗ

В дополнение к сказанному выше, важно отметить, что искусственный интеллект является не только объектом киберугроз, но и фактором формирования новых киберрисков и инструментом атакующих. Генеративные модели и ИИ-агенты могут использоваться злоумышленниками для автоматизации поиска уязвимостей, фишинга, создания вредоносного кода, дипфейков и других атак, одновременно оставаясь эффективным инструментом защиты. Таким образом, ИИ представляет собой технологию двойного назначения, усиливающую возможности как защитников, так и атакующих.

Международные исследования показывают, что бизнес уже осознает этот вызов: 87% организаций относят риски, связанные с ИИ, к числу наиболее быстро растущих киберугроз. Аналогичная тенденция наблюдается и в России: более двух третей компаний указывают на риск использования ИИ для автоматизации разведки и атак, а 40% компаний отмечают, что ИИ-агенты создают не только новые возможности для автоматизации бизнес-процессов, но и становятся самостоятельным источником киберрисков.

Наиболее актуальные угрозы, создаваемые с помощью ИИ, % от ответов



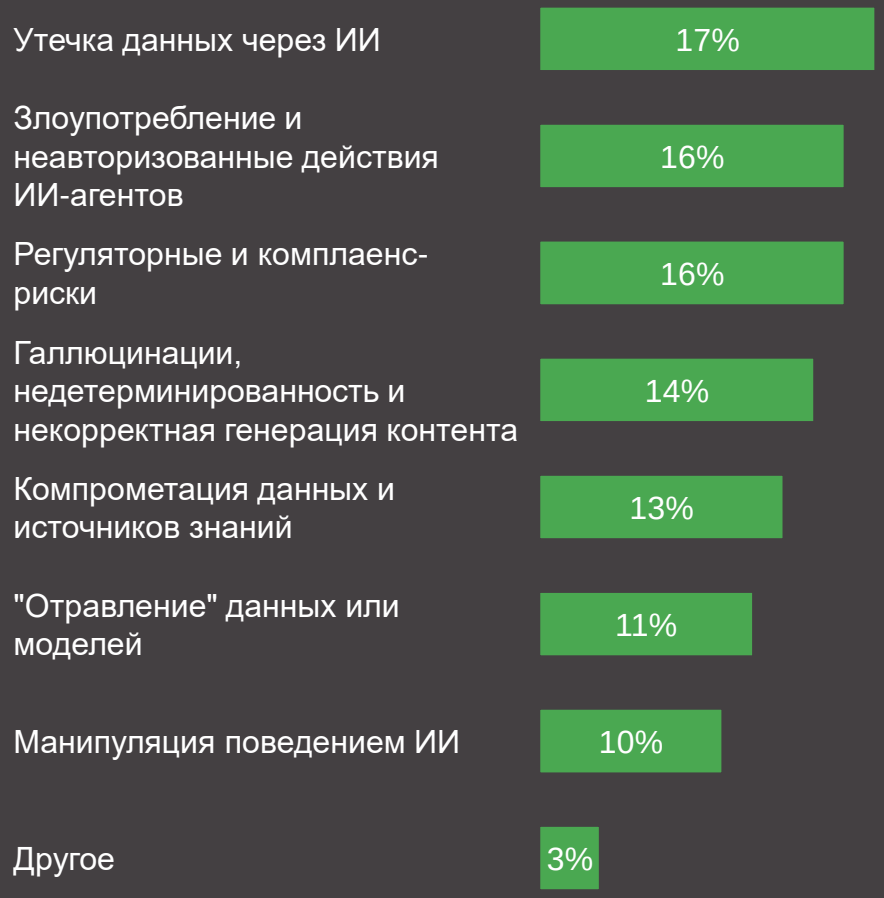
Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

ОТРАСЛЕВОЙ ВЗГЛЯД: БАНКИ, ФИНАНСОВЫЕ УСЛУГИ, СТРАХОВАНИЕ (ФИНТЕХ) (1/3)
Применение ИИ, риски

- Финансовый сектор относится к числу наиболее продвинутых отраслей по уровню стратегического восприятия ИИ.
- Финтех активнее среднего использует ИИ в клиентских сервисах, маркетинге и персонализации, разработке ПО, внутренних помощниках, в процессах ИБ и как часть клиентских продуктов. Понимая значимость ИИ для бизнеса, отрасль находится на этапе перехода от формальных требований к практическому управлению ИИ-рисками.
- В отличие от многих отраслей, где фокус чаще остается на базовых рисках утечки данных, финтех сильнее акцентирует риски, связанные с поведением ИИ: манипуляцией через входные данные, неавторизованными действиями ИИ-агентов, регуляторными и комплаенс-рисками. Риск утечки данных через ИИ остается главным, но его значимость оценивается несколько ниже среднего по всей выборке, что может свидетельствовать о более зрелых механизмах контроля данных.
- Опрошенные компании отрасли демонстрируют уровень контроля использования ИИ сотрудниками выше среднего.

Наиболее критичные риски для ИИ-систем,
% от общих ответов



- Доля компаний отрасли, которые считают ИИ стратегическим приоритетом или критически важной технологией для ключевых процессов **на 9 п.п. выше**, чем в среднем у опрошенных в рамках исследования компаний.
- Уровень финтеха по использованию больших языковых моделей, ИИ-агентов и систем ИИ с дополненным поиском по базам знаний находится **на среднем уровне или даже превышает его**.
- Чаще среднего используются защищенные схемы развертывания ИИ: **44%** респондентов - в частном контуре, и такая же доля использует гибрид частного и публичного контура.
- **50%** опрошенных полностью или скорее удовлетворены защитой ИИ (по сравнению с 41% в среднем).
- Учитывая большую значимость, доля организаций, выделивших бюджет на защиту ИИ, на **14 п.п.** выше, чем в среднем по выборке.

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

ОТРАСЛЕВОЙ ВЗГЛЯД: БАНКИ, ФИНАНСОВЫЕ УСЛУГИ, СТРАХОВАНИЕ (ФИНТЕХ) (2/3)
Меры защиты ИИ

Топ-11 мер защиты ИИ, которые уже реализуются в организациях,
% от опрошенных организаций

- Профиль рисков, связанных с ИИ, отражается в мерах защиты: компании отрасли чаще внедряют не точечные ограничения, а полноценный операционный контур контроля – защиту данных, журналирование и аудит, фильтрацию запросов, непрерывный мониторинг, интеграцию событий ИИ в SOC, контроль автономных решений, разворачивание ИИ в частном контуре. По мере роста автономности ИИ потребности будут и далее смещаться к инструментам постоянного мониторинга, аудита и контроля поведения ИИ в реальных бизнес-процессах.
- Среди топ-5 приоритетов защиты ИИ на ближайшее время:
  - Защита ИИ-агентов — **наиболее часто упоминаемый приоритет среди всех исследованных отраслей** (**56%** респондентов)
  - Формирование стратегии и политик защиты ИИ
  - Защита данных и знаний
  - Аудит и тестирование ИИ
  - Использование ИИ в закрытом контуре
- 25% опрошенных финтех-компаний ожидают существенный рост инвестиций в киберзащиту ИИ (против 9% в среднем) в ближайшие 1-2 года. 25% уже применяют специализированные внешние ИБ-решения (против 16% в среднем). При этом 2/3 опрошенных используют внутренние разработки для защиты ИИ. Отрасль строит защиту сама, тестирует внешние решения и обеспечивает рост бюджета на это направление. Это позволяет рассматривать финтех как одну из наиболее зрелых индустрий с точки зрения формирования спроса на специализированные решения для защиты ИИ.

Источник: результаты анкетирования

ОТРАСЛЕВОЙ ВЗГЛЯД: БАНКИ, ФИНАНСОВЫЕ УСЛУГИ, СТРАХОВАНИЕ (ФИНТЕХ) (3/3)
Мнения экспертов рынка

«Финансовый сектор одним из первых начал массово внедрять искусственный интеллект в процессы, связанные с принятием решений, обработкой клиентских данных, антифрод-системами и автоматизацией сервисов. При этом мы видим, что развитие ИИ происходит значительно быстрее, чем формирование единых подходов к его безопасности.

Для финтеха это особенно чувствительно: любая ошибка модели, утечка данных, манипуляция алгоритмами или недостаточная прозрачность решений напрямую влияет на деньги клиентов, доверие пользователей и соответствие регуляторным требованиям».

Александр Товстолип
Руководитель управления информационной безопасности, Ассоциация ФинТех

«Финтех демонстрирует переход от экспериментального использования ИИ к его промышленному внедрению в ключевые бизнес-процессы. Для отрасли ИИ становится не отдельной технологической инициативой, а элементом операционной архитектуры, влияющим на клиентские сервисы, разработку, антифрод, обработку данных и принятие решений. Это формирует более высокий уровень требований к надежности, прозрачности и управляемости ИИ-систем.

С точки зрения информационной безопасности зрелость финтеха проявляется в том, что управление ИИ-рисками начинает рассматриваться как часть общего контура киберустойчивости.

На первый план выходят не только защита данных и контроль доступа, но и мониторинг поведения ИИ-агентов, аудит решений, контроль автономных действий и интеграция ИИ-событий в процессы SOC и комплаенса.

В перспективе именно финтех может стать ориентиром для других отраслей в формировании практик безопасного внедрения ИИ. Рынок будет смещаться от отдельных защитных решений к комплексным платформам управления ИИ-рисками, где безопасность становится не ограничителем инноваций, а необходимым условием масштабирования ИИ в критически важных процессах».

Сергей Демидов
Директор департамента операционных рисков, информационной безопасности непрерывности бизнеса, Московская биржа

ОТРАСЛЕВОЙ ВЗГЛЯД: ПРОМЫШЛЕННОСТЬ И ПРОИЗВОДСТВО (1/2)
Риски, применение ИИ и мнение эксперта рынка

«С точки зрения промышленной компании существуют два разнонаправленных вектора: защита самого ИИ и применение ИИ для защиты инфраструктуры. По первому — отрасль пока опирается на OWASP LLM Top 10, модель угроз от Яндекса и т. д., но трактует угрозы фрагментарно. В настоящий момент стандарты и практические подходы находятся в процессе становления. Точечные средства вроде LLM-файрволов еще проходят проверку боем против динамично мутирующих атак, и их реальный потенциал еще предстоит подтвердить.

Уже сейчас можно уверенно говорить о том, что обеспечение безопасности невозможно без плотного вовлечения разработчиков ИИ. Для внедрения ИИ в технологические процессы компании важным является применение принципа Secure By Design, который мы используем при разработке ИТ-систем. Именно поэтому очень важно подходить ответственно к форсированному внедрению ИИ в технологические процессы в крупных компаниях. Речь идет не о том, чтобы свернуть все проекты до тех пор, пока мировой кибербез дорастет до нужного уровня зрелости в области защиты ИИ».

Алексей Мартынцев

Директор Департамента защиты информации и ИТ-инфраструктуры, Норникель

Наиболее критичные риски для ИИ-систем,
% от общих ответов

- Доля компаний отрасли, которые считают ИИ стратегическим приоритетом или критически важной технологией для ключевых процессов, составляет **9%** (меньше, чем средний показатель 35%). Доля организаций, находящихся на стадии пилотных проектов, достигает **36%** против 16% в среднем.
- Основные сценарии применения: внутренние помощники и поиск по базам знаний; компьютерное зрение и видеоаналитика; автоматизация обработки документов. Отрасль лидирует по использованию классического ML. ИИ-агенты (36% против 46%) и системы ИИ с дополнительным поиском по базам знаний распространены слабее.
- Доля тех организаций, кто не удовлетворен защитой ИИ и где не выделен бюджет на защиту ИИ, превышает средние показатели по выборке, что свидетельствует о меньшем фокусе внимания на данную область внедрения ИИ.

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

ОТРАСЛЕВОЙ ВЗГЛЯД: ПРОМЫШЛЕННОСТЬ И ПРОИЗВОДСТВО (2/2)
Меры защиты ИИ

Топ-11 мер защиты ИИ, которые уже реализуются в организациях,
% от опрошенных организаций

Топ-5 приоритетов защиты ИИ на ближайший год:

1. Формирование стратегии и политик защиты ИИ
2. Защита данных и знаний
3. Контроль «теневого использования» ИИ
4. Использование ИИ в закрытом контуре
5. Обучение команд

Это отражает переход отрасли от стадии экспериментов к системному управлению ИИ

« Все большую важность приобретает контроль использования ИИ-моделей сотрудниками компании в рабочих целях. Мы стараемся анализировать, в чем состоит запрос сотрудников на использование ИИ, и учитывать это при разработке внутренних моделей, чтобы закрыть имеющиеся потребности безопасным способом. В том числе разрабатываем соответствующие внутренние правила и требования, доносим их до сотрудников. Появляются возможности использования ИИ для решения отдельных производственных задач при условии тщательного контроля со стороны функции ИБ».

Алексей Мартынцев

Директор Департамента защиты информации и ИТ-инфраструктуры, Норникель

Источник: результаты анкетирования

ОТРАСЛЕВОЙ ВЗГЛЯД: РИТЕЙЛ И E-COMM (1/2)
Риски, применение ИИ

- Несмотря на то что доля компаний, считающих ИИ стратегическим или критически важным, находится чуть ниже среднего, отрасль демонстрирует один из самых высоких уровней практического внедрения ИИ. Компании отрасли более активно, чем другие респонденты, используют ИИ в: клиентских чат-ботах и контакт-центрах (100% против 60% в среднем); корпоративном поиске и внутренних помощниках; аналитике и прогнозировании; разработке ПО; ИИ-продуктах для клиентов.
- Уровень защиты ИИ пока отстает от темпов его внедрения. Характерно, что большинство компаний скорее не удовлетворены текущим уровнем защиты ИИ и признают необходимость развития специализированных мер безопасности.
- Ритейл и e-comm характеризуются наиболее выраженной по сравнению со средним концентрацией рисков, связанных с данными (утечки данных, компрометация данных, отравление данных), корректностью работы ИИ и защитой ИИ-агентов.
- Более актуальным, чем для других отраслей, является вопрос неконтролируемого теневого использования ИИ сотрудниками.

Наиболее критичные риски для ИИ-систем,
% от общих ответов

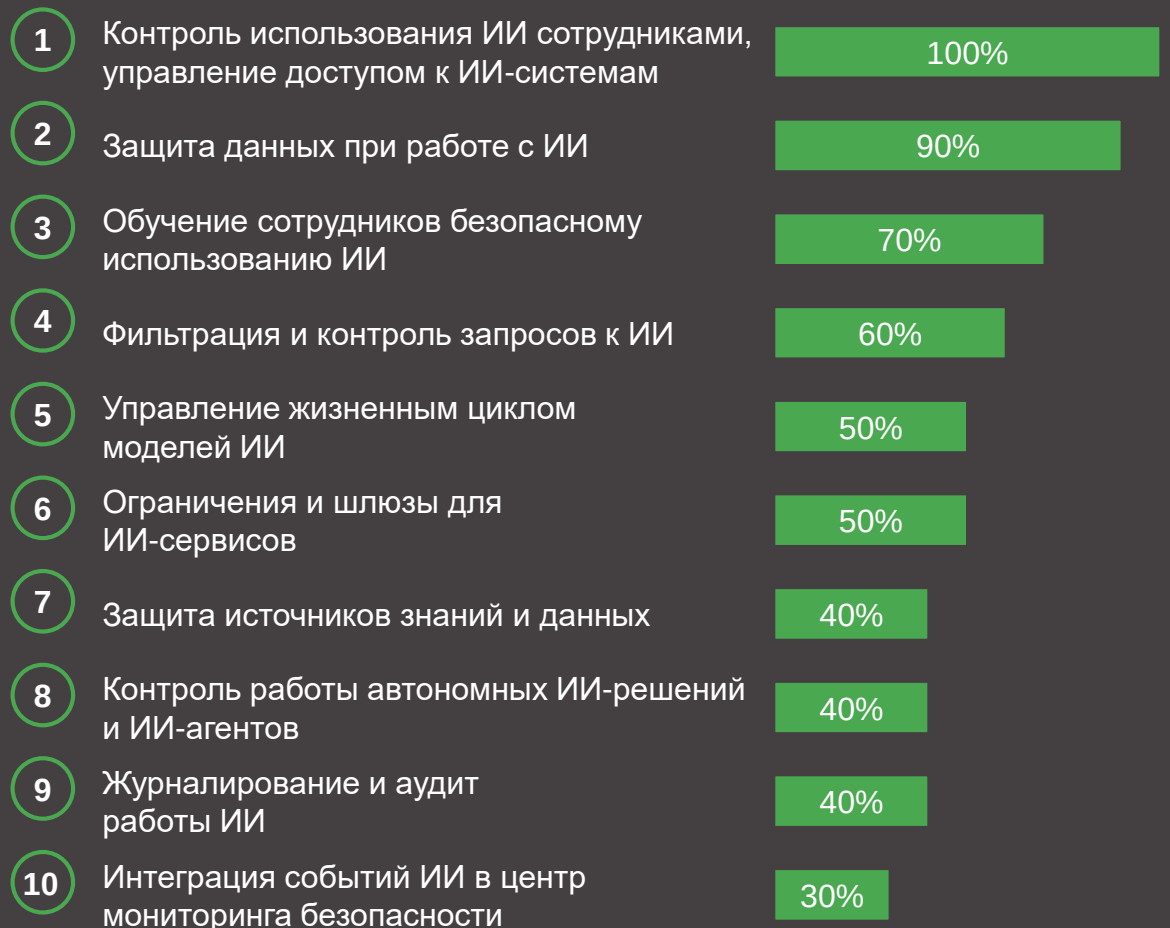
- Доля компаний отрасли, которые считают ИИ стратегическим приоритетом или критически важной технологией для ключевых процессов, составляет **33%** (по сравнению с 35% в среднем). Основной фокус – использование ИИ в отдельных процессах (более половины опрошенных).
- Компании отрасли более активно применяют ИИ-агентов, системы ИИ с дополнительным поиском по базам знаний, дообученные ИИ-модели.
- Отрасль чаще среднего разворачивает ИИ в закрытом контуре, что снижает часть рисков, связанных с внешними ИИ-сервисами, но повышает требования к качеству данных и защите инфраструктуры ИИ.
- **33%** опрошенных полностью или скорее удовлетворены защитой ИИ (по сравнению с 41% в среднем). Доля компаний, где выделен бюджет на защиту ИИ, меньше, чем в среднем по общей выборке.

Источник: результаты анкетирования

Сумма долей может не равняться 100% из-за округлений

ОТРАСЛЕВОЙ ВЗГЛЯД: РИТЕЙЛ И E-COMM (2/2)
Меры защиты ИИ, мнение эксперта

Топ-10 мер защиты ИИ, которые уже реализуются в организациях,
% от опрошенных организаций



- Руководствуясь профилем рисков, компании отрасли существенно чаще среднего внедряют контроль использования ИИ сотрудниками, защиту данных, фильтрацию запросов к ИИ, изоляцию вычислений, маркировку и идентификацию контента, создаваемого ИИ. Отрасль реже инвестирует в мониторинг и аудит ИИ, SOC-интеграцию, что создает потенциальный разрыв между масштабом использования ИИ и контролем.
- Среди топ-5 приоритетов защиты ИИ на ближайший год:
  1. Использование ИИ в закрытом контуре
  2. Защита данных и знаний
  3. Формирование стратегии и политик защиты ИИ
  4. Контроль теневого использования ИИ
  5. Обучение команд
- Интересно отметить, что защита ИИ-агентов пока не рассматривается как ближайший приоритет на ближайшее время.
- Три четверти респондентов ожидают увеличения инвестиций в защиту ИИ в ближайшие 2 года, большинство из них – на умеренном уровне.

« *Исследование хорошо отражает ситуацию в отрасли. ИИ уже активно используется в ключевых клиентских и операционных процессах – от персонализации и антифрода до ИИ-агентов в поддержке и генерации контента. При этом уровень защищенности самих ИИ-систем часто не поспевает за темпами внедрения. Это общий вызов для рынка* ».

Кирилл Мякишев

Директор по информационной безопасности, Ozon

Источник: результаты анкетирования

ЗАКЛЮЧЕНИЕ

По мере того как ИИ становится базовой технологией цифровой экономики, а генеративные модели и ИИ-агенты начинают встраиваться в ключевые бизнес-процессы, вопросы защиты ИИ переходят из категории задач на будущее в область практической необходимости.

В рамках исследования мы стремились показать, что безопасность искусственного интеллекта перестает быть экспериментальной или нишевой темой и постепенно превращается в новую полноценную технологию рынка информационной безопасности.

Глобальный рынок уже формирует новые классы решений: межсетевые экраны и защитные механизмы для ИИ, защита ИИ-систем во время выполнения, безопасность ИИ-агентов, шлюзы безопасного доступа к ИИ, управление ИИ-рисками и использованием ИИ, средства защиты моделей, данных и ИИ-инфраструктуры. Одновременно развивается рынок специализированных услуг. Как уже говорилось, рынок движется от точечных средств защиты к платформенным подходам.

Проведенный в рамках исследования опрос порядка 60 российских организаций показывает, что российский рынок защиты ИИ пока находится на раннем этапе развития, когда специализированные решения и услуги только формируются. Однако интерес к теме неуклонно растет. Многие компании уже внедряют генеративный ИИ и ИИ-агентов, параллельно формируя процессы управления ИИ-рисками, контроля использования ИИ и специализированные механизмы защиты.

Одновременно становится очевидно, что традиционных средств ИБ уже недостаточно для защиты ИИ-систем, а степень развитости решений и объем инвестиций в данную тематику на российском рынке уступает глобальным практикам. Это создает как вызовы, так и возможности для развития отечественных поставщиков технологий и услуг ИБ.

В перспективе способность безопасно внедрять и использовать ИИ может стать не только вопросом защиты, но и одним из важных факторов устойчивости и конкурентоспособности бизнеса и всей цифровой экономики.

АВТОРЫ ИССЛЕДОВАНИЯ И БЛАГОДАРНОСТЬ РЕСПОНДЕНТАМ И ЭКСПЕРТАМ, УЧАСТВОВАВШИМ В ИССЛЕДОВАНИИ

Б1

Сергей Салов
Партнер Группы компаний Б1
sergey.salov@b1.ru

При участии старшего менеджера
Анны Патраковой и консультанта
Ангелины Прохоровой

ГК «Солар»

Мария Курбатова
Руководитель группы маркетинговой аналитики ГК «Солар»
Василий Мирошкин
Менеджер по работе со стартапами и технологическими партнерами ГК «Солар», руководитель программы CyberStage

При участии директора по венчурным инвестициям и партнерствам ГК «Солар»
Владислава Рассказова

АФТ

Александр Товстолип
Руководитель управления информационной безопасности Ассоциации ФинТех
Сергей Лапин
Эксперт по информационной безопасности Ассоциации ФинТех

HiveTrace

Евгений Кокуйкин
Основатель, генеральный директор HiveTrace
Иван Василенко
Директор по клиентской стратегии

Авторы выражают признательность всем участникам опроса, а также персонально следующим экспертам:

- Сергей Демидов, Московская биржа
- Алексей Мартынцев, Норникель
- Кирилл Мякишев, Ozon
- Денис Султанов, БКС
- Александр Шепилов, Ростелеком

ИНФОРМАЦИЯ ОБ ИССЛЕДОВАНИИ

Информация, содержащаяся в настоящей публикации, предназначена лишь для общего ознакомления, в связи с чем она не может служить основанием для вынесения профессионального суждения. Группа компаний Б1 и другие авторы исследования не несут ответственности за ущерб, причиненный каким-либо лицам в результате действия или отказа от действия на основании сведений, содержащихся в данной публикации. По всем конкретным вопросам следует обращаться к специалисту по соответствующему направлению.

О ГРУППЕ КОМПАНИЙ Б1

Группа компаний Б1 предлагает многопрофильные услуги в сфере аудита, стратегического, технологического и бизнес-консалтинга, сделок, оценки, налогообложения, права и сопровождения бизнеса.

Мы работаем свыше 35 лет в России и более 25 лет в Беларуси. За это время в компаниях группы создана сильная команда специалистов с обширными знаниями и опытом реализации сложнейших проектов. Наша практика представлена в 12 городах: Москве, Минске, Владивостоке, Екатеринбурге, Казани, Краснодаре, Новосибирске, Ростове-на-Дону, Самаре, Санкт-Петербурге, Тольятти и Челябинске.

Группа компаний Б1 помогает клиентам находитьновые решения, расширять, трансформировать и успешно вести свою деятельность, а также повышать свою финансовую устойчивость и кадровый потенциал.

.

b1.ru

О ГРУППЕ КОМПАНИЙ «СОЛАР»

Группа компаний «Солар» — лидирующий провайдер комплексной кибербезопасности. Ключевые направления деятельности — предоставление услуг и сервисов в области информационной безопасности, разработка собственных ИБ-продуктов, обучение ИБ-специалистов, аналитика и исследование киберинцидентов.

С 2015 года предоставляет ИБ-решения организациям от малого бизнеса до крупнейших предприятий ключевых отраслей. Под защитой «Солара» — более 1 500 крупнейших компаний России. Компания работает в направлениях безопасной разработки программного обеспечения, управления доступом, защиты корпоративных данных, детектирования хакерских атак и угроз, что позволяет закрывать максимум потребностей заказчиков.

Группа компаний предлагает сервисы первого и крупнейшего в России коммерческого SOC — Solar JSOC, экосистему управляемых сервисов ИБ — Solar MSS. По данным независимых аналитиков, «Солар» входит в топ-5 европейских и топ-15 мировых сервис-провайдеров по объему бизнеса.

Работа центра исследования киберугроз Solar 4RAYS нацелена на изучение тактик киберпреступников. Полученные аналитические данные обогащают разработки Центра технологий кибербезопасности.

Линейка собственных продуктов включает DLP-решение Solar Dozor, шлюз веб-безопасности Solar webProxy, IdM-систему Solar inRights, PAM-систему Solar SafeInspect, анализатор кода Solar appScreener, решение для автоматизации мониторинга и реагирования Solar SIEM и другие.

rt-solar.ru

ОБ АССОЦИАЦИИ ФИНТЕХ

Ассоциация ФинТех учреждена в конце 2016 г. по инициативе Банка России и ключевых участников финансового рынка как межотраслевая площадка для взаимодействия банков, страховых и ИТ-компаний, профессиональных участников рынка ценных бумаг, телеком-операторов, финтехов и регулятора.

В настоящий момент в состав АФТ входят: Банк России, Сбер, ВТБ, Альфа-Банк, Газпромбанк, Национальная система платежных карт (НСПК), Т-Банк, Райффайзенбанк, Россельхозбанк, ДОМ.РФ, Московская Биржа, Страховой Дом ВСК, ПСБ, Совкомбанк; ассоциированные члены: Ак Барс Банк, РНКО «Платежный центр», МТС, Ростелеком, АльфаСтрахование, Капитал Лайф Страхование Жизни, Московский кредитный банк, Кредит Европа Банк (Россия), Банк Синара, Яндекс, Центр-Инвест, ОТП Банк, «Диасофт», «ФлексСофт», Авито, Т1, РЕД СОФТ, Билайн, БПС, Банк Точка, УК «ДОХОДЪ», РЕСО-Гарантия, Банк ВБРР, МТС Банк.

fintechru.org

О HIVETRACE (ООО «ХАЙВТРЕЙС»)

«Хайвтрейс» - российская компания, специализирующаяся на безопасности искусственного интеллекта и разработке решений для защиты корпоративных ГенИИ-систем. Компания разработала и продолжает развивать и продвигать на рынке продукты, обеспечивающие безопасность внедрения и эксплуатации ИИ, в т.ч.:

HiveTrace AI Firewall - система мониторинга ИИ-систем в реальном времени:

HiveTrace Red - автоматизированный red-teaming для ИИ:

Команда HiveTrace объединяет инженеров и исследователей безопасности ИИ, сотрудничает с ведущими вузами и участвует в развитии отраслевых стандартов, создавая полноценный цикл защиты: тестирование, мониторинг и предотвращение атак.

В штате компании работают эксперты в области Data Science и кибербезопасности с успешным опытом разработки и внедрения ПО в компаниях масштаба крупных компаний. Сотрудники “Хайвтрейс” регулярно участвуют в профильных конференциях, выступают в качестве спикеров в вебинарах, подкастах и дают экспертные комментарии в СМИ.

hivetrace.ru