

Искусственный интеллект в России



Вступительное слово и введение



Павел Охонин

Партнёр KAMA FLOW

«Как компания, активно инвестирующая в разные технологические рынки, KAMA FLOW подходит к анализу каждой отрасли с точки зрения достаточной ёмкости рынка и доступности чека для частного фонда.

Рынок AI — один из исторических инвестиционных приоритетов компании. Однако появление и стремительное развитие generative AI резко поменяли конъюнктуру отрасли, сместив акценты с классического машинного обучения на инфраструктурные и прикладные решения нового поколения.

Задачей настоящего исследования стала оценка актуальной возможности деятельного присутствия частного инвестора в AI. Мы стремимся ответить на вопрос — способствуют ли этому современные реалии или рынок остаётся уделом крупных корпораций?».

«Искусственный интеллект формирует новый технологический и инвестиционный цикл, в котором корпорации, государства и инвесторы одновременно стремятся занять позиции на ещё не сложившемся рынке. При этом масштаб вложений и скорость появления новых решений пока опережают формирование устойчивых бизнес-моделей, что делает рынок особенно сложным для анализа.

В России этот технологический переход происходит в особых условиях. Глобальная гонка за лидерство в AI сочетается с ограниченным доступом к капиталу, вычислительной инфраструктуре и зарубежным технологиям, а также с необходимостью развивать собственные модели, продукты и компетенции.

В результате российский AI-рынок не повторяет глобальную модель, а формирует собственную конфигурацию участников. В данной работе рассматривается российский AI-рынок в момент его становления: что определяет его развитие, где создаётся реальная экономическая ценность и какие возможности открываются для российских компаний и частного капитала».

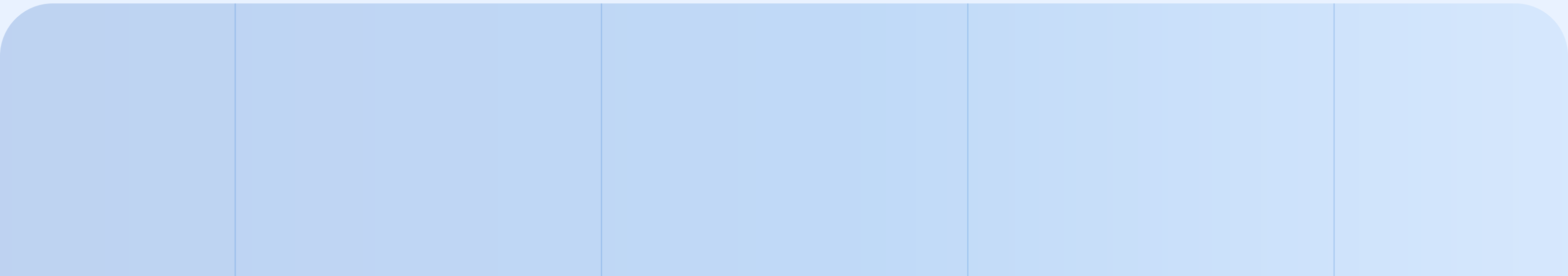


Карина Ручкина

Инвестиционный аналитик
KAMA FLOW

Содержание

01	Фокус исследования	5
02	Капитал в AI: глобальный и российский рынок	7
03	Экономика AI-решений	14
04	Российские AI-стартапы: ограничения и условия роста	23
05	Карта инвестиционных возможностей	27
06	Специфика и барьеры развития российского рынка	33
07	Основные выводы	36



Фокус исследования



Ключевой вопрос KAMA FLOW

Есть ли в российском AI ниши с потенциалом кратного роста для независимых частных компаний?

Глобальный AI-рынок продолжает привлекать значительный объем капитала: только в I квартале 2026 года венчурные инвестиции в сектор составили \$242 млрд, при этом крупнейшие мега-раунды сосредоточены вокруг разработчиков фундаментальных моделей.

Одновременно ведущие технологические компании расширяют инфраструктурные программы, включая строительство вычислительных кластеров, развитие дата-центров и инвестиции в производство специализированных AI-чипов.

В России многомиллиардные инвестиции остаются преимущественно уделом корпораций и недоступны частному инвестору, чей средний чек редко превышает 1 млрд ₽.

В этих условиях для инвестора важно понять, где за пределами крупных экосистем смогут появиться самостоятельные компании, способные превращать AI-решения в масштабируемый бизнес с понятной экономикой, возможностью роста и реалистичным сценарием выхода?

Критерии «идеального кандидата»

Масштаб

3–5 млрд ₽ выручки как достижимый объем

Потребность в капитале

Возможность инвестировать до 1 млрд ₽

Exit

Потенциал IPO. При меньшем масштабе бизнеса (до 1 млрд ₽ выручки) — понятная логика выхода через M&A

Доходность

Потенциал IRR >30–40% годовых

Направления исследования

Капитал в AI

Сопоставили динамику венчурных инвестиций и вложения крупнейших технологических корпораций в России и мире.

Экономика AI-решений

Рассмотрели особенности монетизации и структуры затрат AI-продуктов, сравнили расходы при разных вариантах развертывания моделей. Проанализировали цепочку создания ценности AI-решений для заказчика.

Условия роста независимых компаний

Проанализировали ограничения развития ИИ-стартапов и их конкуренцию с внутренней разработкой компаний.

Карта инвестиционных возможностей

Сегментировали рынок, оценили потенциал отдельных направлений к 2030 году и сопоставили их по инвестиционной привлекательности и капиталоемкости.

Специфика российского рынка

Рассмотрели различия в темпах внедрения ИИ по отраслям и бизнес-функциям, а также инфраструктурные, регуляторные и организационные барьеры.

В рамках исследования команда Kama Flow провела серию интервью с представителями IT-корпораций, AI-стартапов и промышленных компаний, чтобы сопоставить взгляды на российский рынок ИИ с позиций поставщиков технологий и инфраструктуры, разработчиков решений и корпоративных заказчиков.

Цель исследования

Определить наиболее привлекательные рыночные сегменты для вложений в независимых AI-игроков с потенциалом масштабирования и понятной моделью доходности для частного инвестора

Генеративный AI формирует **новый класс B2B-продуктов**

Если классический ML давно стал зрелой частью корпоративной инфраструктуры и используется в бизнесе как инструмент анализа данных, то генеративный AI позволяет автоматизировать интеллектуальный труд и конкретные бизнес-процессы.

Классический ML



Алгоритмы, решающие задачи на выявление паттернов:

Предиктивная аналитика

Поиск аномалий

Кластеризация

Классификация



Генеративный AI



Решения на основе базовых моделей (LLM/мультимодальные):

Позволяет работать с неструктурированными данными (текст, изображение, звук)

Требует много вычислительных ресурсов для обучения (GPU)

Генеративный AI автоматизирует рутинные бизнес-процессы и более сложные задачи, которые раньше требовали высокой квалификации специалистов и значительных затрат на их труд, – например, разработку программного кода. Одновременно формируются новые продуктовые сегменты – копилоты, AI-агенты и RAG-системы.



Эволюция AI



1990-е

Банковские скоринговые модели



2000-е

Рекомендательные системы товаров и контента



2010-е

Развитие компьютерного зрения



2017

Изобретение архитектуры Transformer



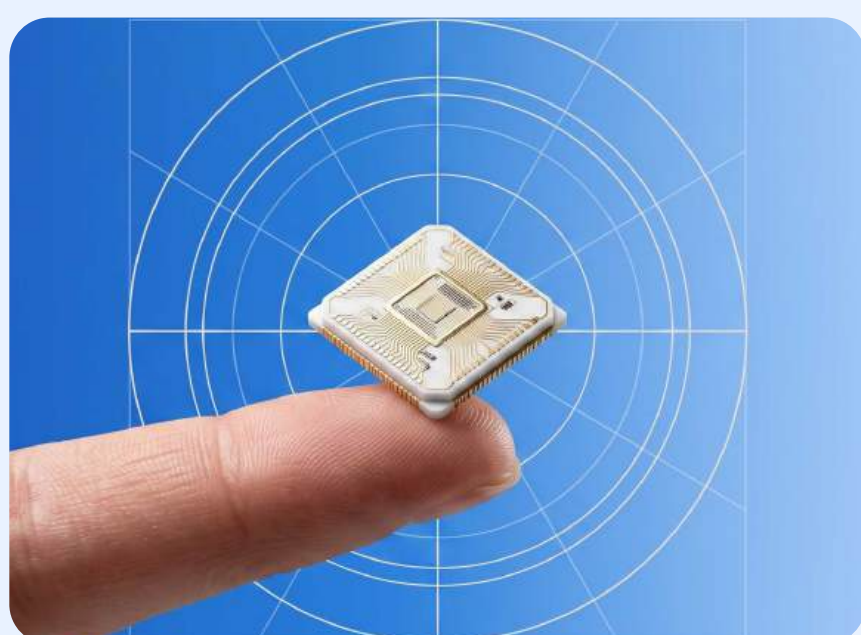
2023

ChatGPT сделал генеративный AI массовой технологией



2025+

Развитие агентных сценариев: GenAI теперь выполняет многошаговые задачи

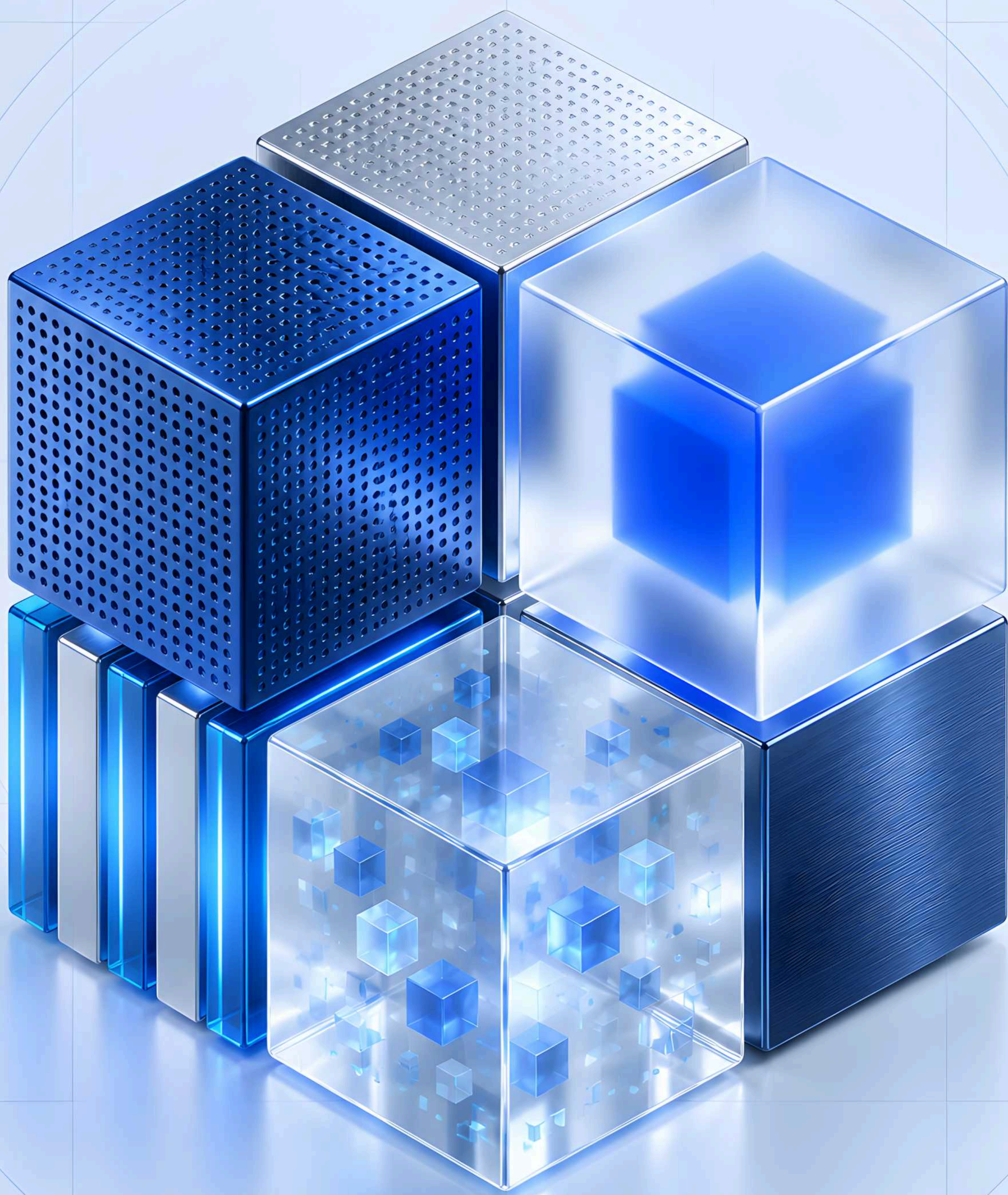


В фокусе исследования – именно решения на основе генеративного AI

Рынок находится в стадии формирования, продуктовая экономика еще не устоялась, а ключевым фактором успеха для новых компаний станет способность превратить технологию в масштабируемый бизнес.

Капитал в AI:

глобальный и российский рынок



Капитал в AI: глобальный и российский рынок

AI обеспечил рекорд глобального VC-рынка в I полугодии 2026 года

\$386

млрд инвестиций в AI

76%

доля AI во всём VC-рынке

\$510

млрд – весь VC-рынок

AI меняет структуру глобального венчурного рынка

2021 — I полугодие 2026 · \$ МЛРД

За шесть месяцев инвестиции в AI почти вдвое превысили объём финансирования AI-стартапов за весь 2025 год.

≈2x

объём инвестиций за 6 месяцев к уровню всего 2025 года

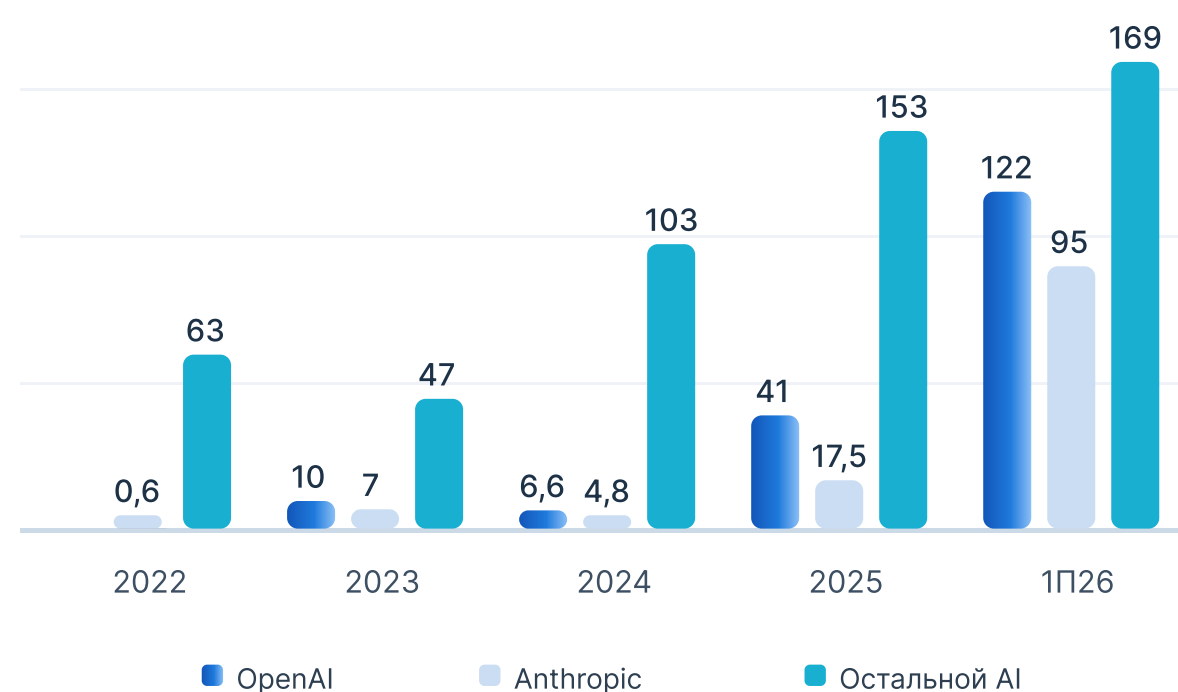


Компании, занимающиеся разработкой моделей, привлекли в I квартале 2026 года \$178 млрд – 74% от VC-инвестиций в AI



\$178 млрд в 1К2026 · 74% AI-инвестиций

OpenAI и Anthropic привлекли \$217 млрд – 56% AI-инвестиций и 43% всего VC-рынка за 1П2026



\$217 млрд · 56% AI-инвестиций · 43% всего VC

Источники:

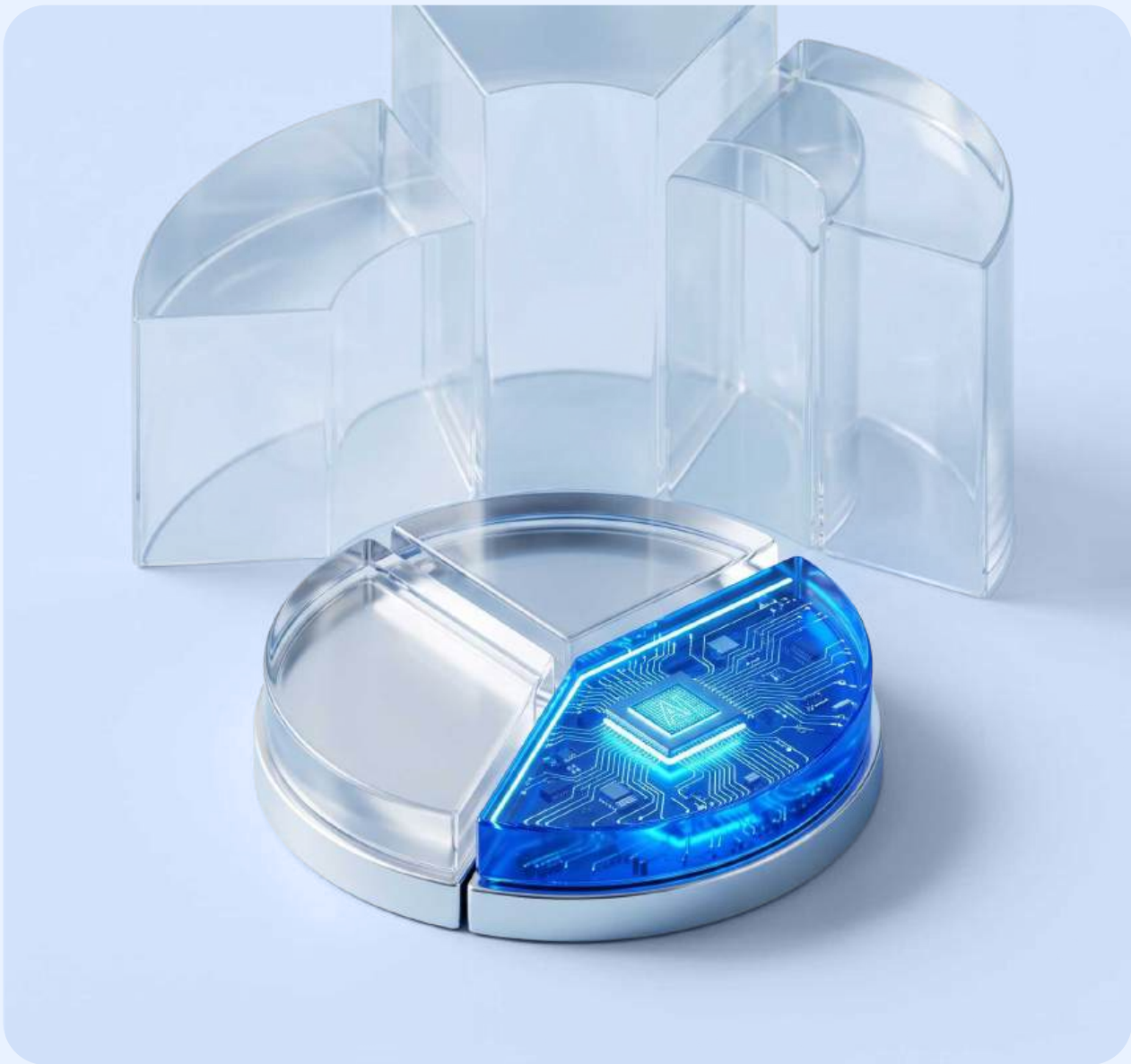
Crunchbase News – Q1 2026 Shatters Venture Funding Records As AI Boom Pushes Startup Investment To \$300B

Crunchbase News – 6 Charts That Show The Big AI Funding Trends Of 2025

Crunchbase News – Global Startup Investment Hit Record \$510B In H1 2026 As AI Boom Accelerates Funding And Exits

Российский венчурный рынок в 2025 году составил \$159 млн, из них около \$51 млн пришлось на AI-проекты

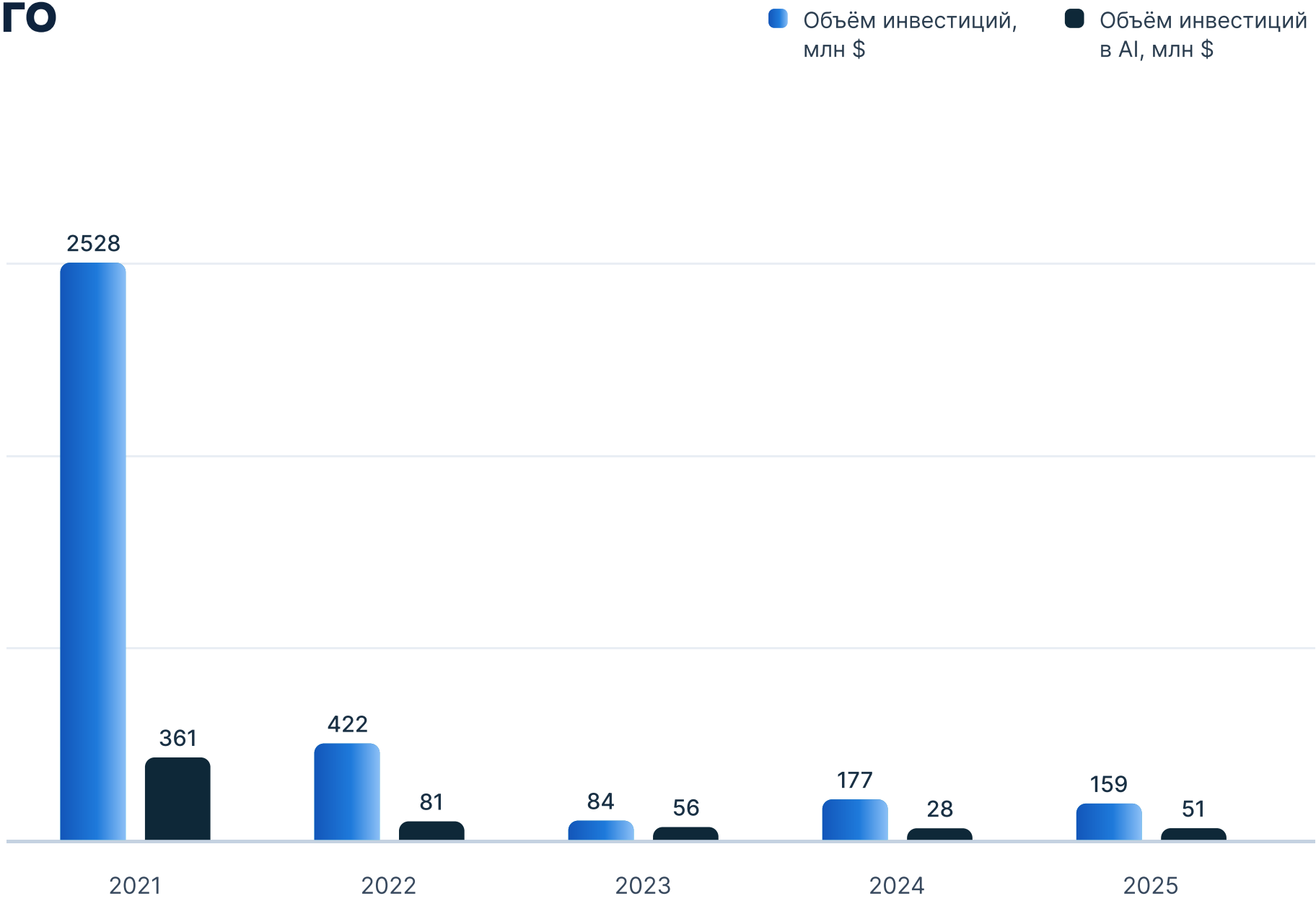
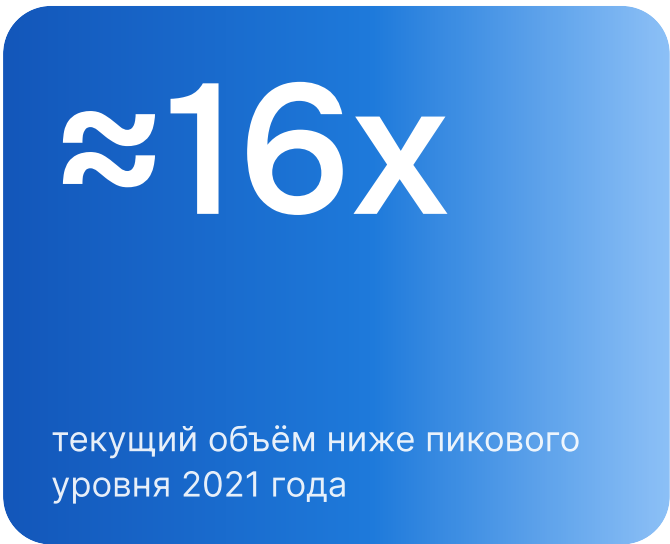
Для сравнения: в 2021 году объём российского венчурного рынка достигал \$2,5 млрд — примерно в 16 раз больше, чем в 2025-м. Сокращение рынка связано в том числе с уходом иностранного капитала после 2022 года.



Динамика российского венчурного рынка

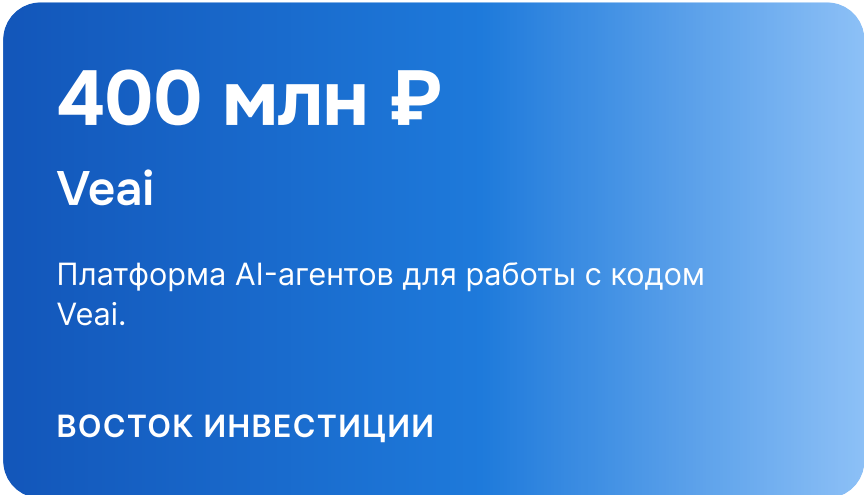
2021 — 2025 · \$ млн

Инвестиционный интерес к AI в России присутствует, однако по темпам развития российский рынок значительно отстает от глобального.



Крупнейшие AI-сделки 2025 года

В выборку вошли сделки с компаниями, использующими AI в продуктах, в том числе в случаях, когда он не является ключевой технологией.



На фоне ограниченного венчурного рынка ключевым драйвером развития AI в России становятся инвестиции крупнейших корпораций

Российский AI-рынок развивается преимущественно по **корпоративной модели** – основной объём инвестиций сосредоточен внутри крупнейших экосистем, тогда как венчурное финансирование независимых AI-стартапов остаётся ограниченным.

150 млрд ₽

– инвестиции Сбера в AI в 2025 году. Общий объём трёхлетней инвестиционной программы – 600 млрд ₽.

55 млрд ₽

инвестировал Яндекс в 2025 году в расширение команды, развитие процессов разметки данных и наращивание вычислительных мощностей.

ИТ-КОРПОРАЦИИ · 2025

220 млрд ₽

Объём публично заявленных инвестиций российских ИТ-корпораций в AI в 2025 году.

Ключевыми игроками российского AI-рынка являются Сбер и Яндекс

Их преимущество заключается не только в наличии собственных моделей, но и в контроле над значительной частью AI value chain – от вычислительной инфраструктуры и разработки фундаментальных моделей до прикладных сервисов и массовой дистрибуции через пользовательские сценарии собственных экосистем. Доступ к массовой аудитории позволяет компаниям масштабировать новые AI-продукты без значительных затрат на привлечение клиентов и накапливать данные для дальнейшего совершенствования моделей.

Слой AI Value Chain	Сбер	Яндекс
• Вычислительная инфраструктура	Партнёрские мощности Cloud.ru	Yandex Cloud
• Фундаментальные модели	GigaChat	YandexGPT и Alice AI
• Платформа для разработки	ГигаЧат Бизнес	Yandex AI Studio
• Прикладные решения	СберМедИИ и др.	Алиса AI
• Дистрибуция	Банковская экосистема, умные устройства	Поиск, Браузер, умные устройства

В отличие от глобального рынка, где ключевую роль в разработке фундаментальных моделей играют независимые AI-компании (в частности, OpenAI и Anthropic), в России разработка моделей полностью сосредоточена внутри BigTech-корпораций.

«В России есть две компании – «Яндекс» и «Сбер», которые могут себе позволить участвовать в этой гонке [фундаментальных моделей – прим. KAMA FLOW]. В эту гонку больших бюджетов и суперкоманд входить точно уже поздно – возможностей для этого для какой-то третьей компании в России я не вижу».

Александр Ведяхин
Первый заместитель председателя правления
СберБанка
Интервью Forbes, 24 февраля 2026 года

«Разработка моделей, их обучение, эксперименты с моделями – это очень дорогостоящее занятие. Конечно, позволить его могут себе только достаточно крупные компании и те компании, которые научились зарабатывать на искусственном интеллекте».

Герман Греф
Президент, председатель правления СберБанка
Комментарий «Интерфаксу», 19 ноября 2025 года

Доступ к китайскому open source создаёт временное окно возможностей, но оно может закрыться. Поэтому собственные модели отечественных вендоров приобретают стратегическое значение для AI-рынка.

«Стратегия многих российских AI-директоров сегодня строится на том, чтобы покупать GPU и разворачивать китайские open-source-модели в расчёте, что со временем они достигнут уровня Claude.

Но по мере роста качества китайские вендоры уже начинают закрывать коммерческое использование моделей и перестают публиковать их веса: открытыми, скорее всего, останутся только решения, которые не дают стратегического преимущества.

Для Китая наиболее качественные модели становятся критической технологией, поэтому рассчитывать на постоянный доступ к новым поколениям open source нельзя. Российскому рынку важно успеть локализовать наиболее сильные из доступных сейчас китайских моделей, поскольку в дальнейшем такой возможности может уже не быть».

Армен Амирханян
Директор по развитию искусственного интеллекта
Московской биржи

«Я думаю, что мы, как страна, можем катастрофически отстать, если мы откажемся от использования открытых моделей. <...> И ни в коем случае нельзя противопоставлять одно второму. Мы просто упустим свой шанс в большую эпоху технологической трансформации человечества».

Тигран Худавердян
Член правления АО «Яндекс рус»
Комментарий ТАСС, 2 сентября 2026 года

Венчурный AI-рынок и вложения корпораций взаимосвязаны. AI-стартапы — источник спроса на инфраструктуру гиперскейлеров: их модели, облака, GPU-мощности и дата-центры

По данным за 2025 год:



Стартапы формируют спрос на вычисления – гиперскейлеры монетизируют этот спрос через инфраструктуру

По оценкам на середину 2026 года, капитальные расходы пяти крупнейших гиперскейлеров (Microsoft, Alphabet, Amazon, Meta, Oracle) прогнозируются на уровне \$820 млрд – по своему масштабу эти затраты уже обходят прогноз расходов CapEx нефтегазовой отрасли (\$620 млрд). Основной драйвер роста CapEx гиперскейлеров – инвестиции в развитие инфраструктуры для AI. Большой объём затрат ожидается только у Министерства обороны США (\$1010 млрд). Более того, глобальные инвестиции в AI уже превосходят объём вложений в такие мегапроекты XX века, как Манхэттенский проект (проект США по разработке ядерного оружия) и космическую программу Аполлон.

Так, в 2025 году объём венчурных инвестиций в AI составил около \$212 млрд, при этом CapEx гиперскейлеров – порядка \$450 млрд. Можно заметить, что на каждый \$1 венчурных инвестиций в AI в 2025 году приходилось примерно \$2 CapEx гиперскейлеров. Подобная тенденция ожидается и в 2026 году.

Венчурный капитал финансирует разработку моделей и прикладных продуктов, увеличивая спрос на вычислительные мощности гиперскейлеров. Гиперскейлеры, в свою очередь, расширяют инфраструктуру и выступают драйверами рынка – масштабируя мощности, они стимулируют дальнейший рост числа AI-стартапов.

Таким образом, гиперскейлеры не только монетизируют растущий спрос на вычисления, но и сами создают условия для дальнейшего расширения AI-рынка.

Масштаб инфраструктурных инвестиций в AI уже сопоставим с крупнейшими капиталоемкими отраслями и государственными программами

Прогноз расходов на 2026 год и стоимость отдельных программ, млрд \$*



* с учётом инфляции, в ценах 2026 года
** По оценке The Planetary Society, в ценах 2025 года

Ключевая структурная проблема российского рынка — недостаток венчурного финансирования, который ограничивает формирование внешнего спроса на коммерческую AI-инфраструктуру

Крупнейшими поставщиками AI-инфраструктуры в России являются ведущие IT-корпорации — уже упомянутые Сбер и Яндекс, а также Selectel, MTC, VK и другие игроки.

52x

корпоративных инвестиций на каждый \$1 VC

Таким образом, на каждый \$1 венчурного финансирования AI в России приходится около \$52 корпоративных инвестиций – против соотношения 1 к 2 на глобальном рынке.



МИР · 2025

CapEx гиперскейлеров Big 5

\$449 млрд

Глобальный венчурный AI-рынок

\$212 млрд

Big 5: Microsoft, Alphabet, Amazon, Meta, Oracle

~2x

На каждый доллар VC-инвестиций в AI — два доллара CapEx гиперскейлеров

РОССИЯ · 2025

Инвестиции BigTech в AI

\$2,6 млрд*

Российский венчурный AI-рынок

\$51 млн

*Инвестиции крупнейших поставщиков инфраструктуры — Сбер и Яндекс

~52x

На каждый доллар VC-инвестиций в AI — 52 доллара инвестиций BigTech в России

В отличие от глобального рынка, где стартапы создают спрос на инфраструктуру гиперскейлеров, в России спрос концентрируется внутри BigTech и крупных корпораций.

Недостаток венчурного финансирования сдерживает развитие независимых AI-решений и ограничивает формирование широкого спроса на вычислительную инфраструктуру.

ПУБЛИЧНО ЗАЯВЛЕННЫЕ ИНВЕСТИЦИИ · 2025

220 млрд ₽

Согласно публичным заявлениям, совокупный объём инвестиций этих компаний в AI в 2025 году, как отмечалось ранее, составил около \$2,6 млрд, или 220 млрд рублей. Преимущественно эти средства были направлены на развитие вычислительной инфраструктуры, но также в этот объём включены инвестиции в фундаментальные модели, AI-продукты и другие направления.

Фактический объём вложений может быть существенно выше — расчёт учитывает только инвестиции, которые компании публично отнесли непосредственно к AI, тогда как часть расходов на AI-инфраструктуру может быть включена в более крупные статьи капитальных и операционных затрат.

Если провести аналогию с международным рынком и сравнить объём венчурных инвестиций в России с вложениями российских крупнейших IT-корпораций, то мы обнаружим структурный дисбаланс.

В 2025 году венчурные инвестиции в российские AI-компании составили около \$51 млн, тогда как вложения крупнейших технологических корпораций достигли примерно \$2,65 млрд.

«На Западе новые игроки могут быстро масштабироваться благодаря развитому рынку и большому объёму доступного капитала.

Компания может сначала создать продукт, привлечь инвестиции ещё до появления выручки, а затем сформировать команду продаж и выйти на рынок. В России такая модель практически не работает: даже при наличии качественного продукта привлечь капитал на go-to-market крайне сложно».

Алексей Постернак
Директор по инвестициям и развитию MWS

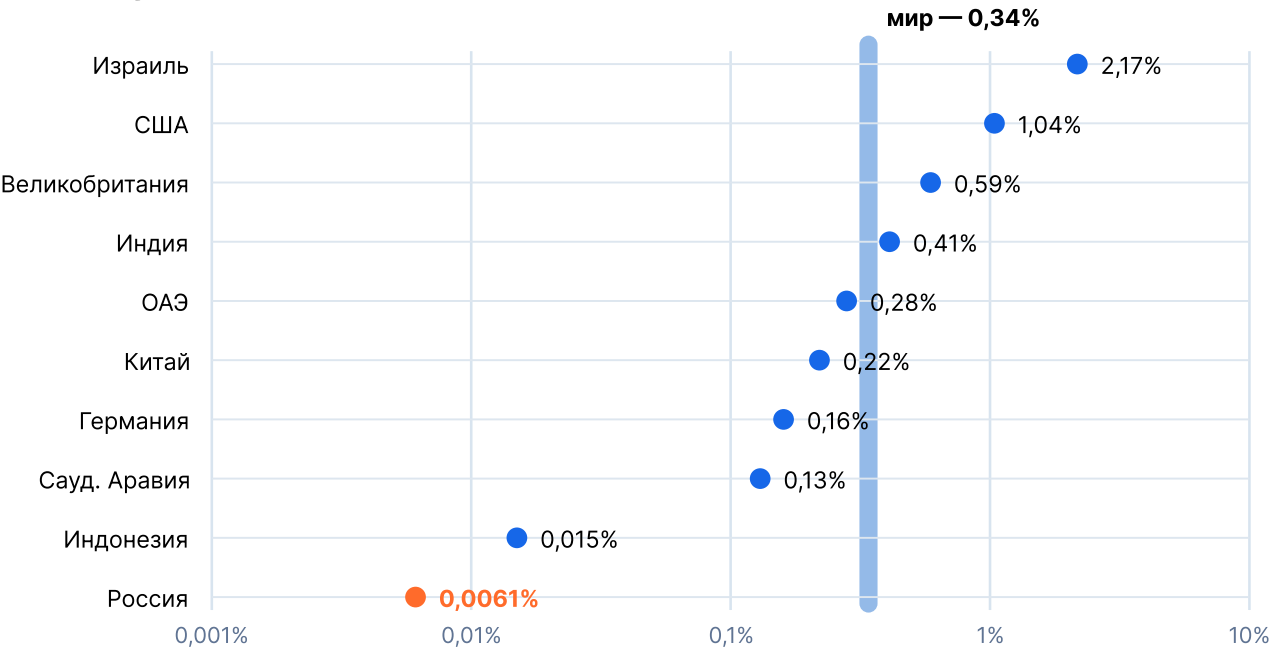


84 млрд ₽ дополнительного венчурного капитала ежегодно необходимо для развития независимого AI-рынка в России

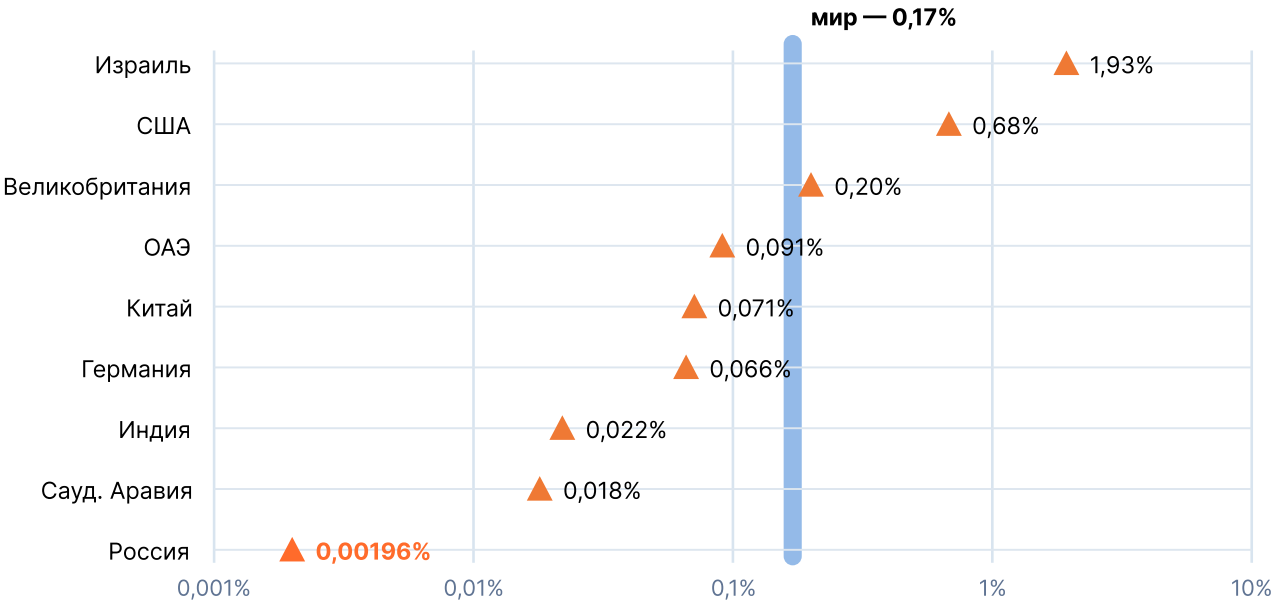
По оценке KAMA FLOW, российскому рынку требуется 84 млрд ₽ дополнительных частных венчурных инвестиций ежегодно для формирования устойчивого пула независимых AI-компаний.

Их масштабирование будет стимулировать спрос на инфраструктуру российских провайдеров: ЦОД, публичные и частные облака, фундаментальные модели и платформенные решения для развертывания AI-продуктов

Венчурные инвестиции, % ВВП, 2025



Венчурные инвестиции в AI, % ВВП, 2025



Россия отстаёт от мирового уровня финансирования AI

Проблема заключается в крайне малом абсолютном объёме частного капитала. В 2025 году общий объём венчурных инвестиций в России составил 159 млн \$, или 0,0061% ВВП.

Из них около 51 млн \$ было инвестировано в компании с AI-составляющей – 32% всего венчурного рынка или 0,002% ВВП.

Для сравнения, мировой объём венчурных инвестиций в 2025 году достиг 405 млрд \$, или 0,34% мирового ВВП. Вложения в AI-компании составили 212 млрд \$, или 0,17% мирового ВВП.

По объёму венчурного финансирования относительно ВВП Россия отстаёт от мирового уровня примерно в 56 раз. Если в качестве более консервативного бенчмарка взять Германию – экономику, сопоставимую по своему размеру с российской, – разрыв составит около 26 раз: 0,0061% ВВП в России против 0,16% в Германии.

В сегменте AI отставание ещё значительнее. Венчурное финансирование AI-компаний в России примерно в 87 раз ниже общемирового уровня и в 34 раза ниже уровня Германии. Выход на немецкий уровень потребовал бы около 1,7 млрд \$ дополнительных инвестиций ежегодно, а на общемировой – около 4,4 млрд \$.

По оценке KAMA FLOW, реалистичным ориентиром является привлечение 84 млрд ₽ дополнительного частного венчурного капитала ежегодно (около 1 млрд \$). Дефицит финансирования носит системный характер: на ранних стадиях он связан с низкой активностью бизнес-ангелов, а на последующих – с ограниченным объёмом капитала частных венчурных фондов и слабым развитием корпоративного венчура. Речь идёт не о разовом вливании, а о регулярном потоке капитала по всей цепочке финансирования – от первых раундов на стадиях pre-seed и seed до раундов на стадиях A/B для роста и масштабирования зрелых компаний.

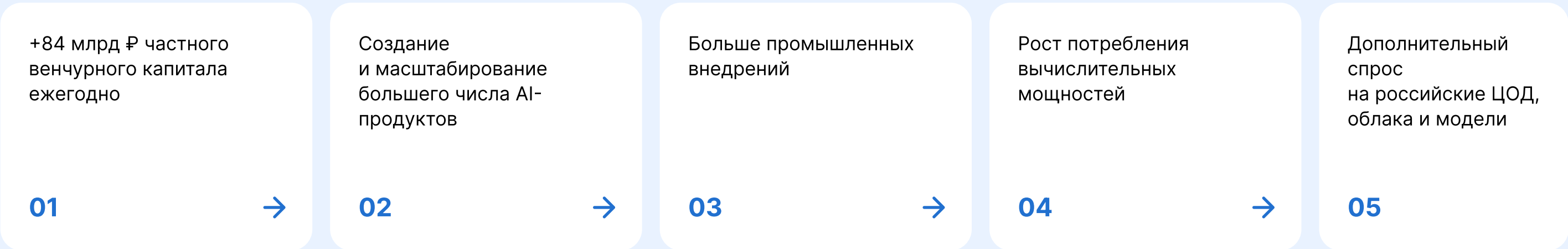
Эффект будет распространяться далеко за пределы самих получателей инвестиций. Масштабирование AI-продуктов увеличит потребление вычислительных мощностей как со стороны разработчиков AI-решений, так и внутри контуров их клиентов. Это создаст дополнительный регулярный спрос на мощности ЦОД, публичные и частные облака, хранение и обработку данных, фундаментальные модели российских вендоров и платформенные решения для развертывания AI-продуктов.

Масштабирование AI-продуктов формирует дополнительный спрос на российскую инфраструктуру

- Мощности ЦОД01
- Облака02
- Модели и платформы03

Таким образом, финансирование независимых AI-компаний становится механизмом развития всего российского IT-рынка, формируя дополнительный спрос на решения крупнейших инфраструктурных провайдеров. К числу прямых бенефициаров такого спроса можно отнести, например, Yandex Cloud, Cloud.ru, MWS, РТК-ЦОД и Selectel, а также разработчиков фундаментальных моделей – Сбер и Яндекс.

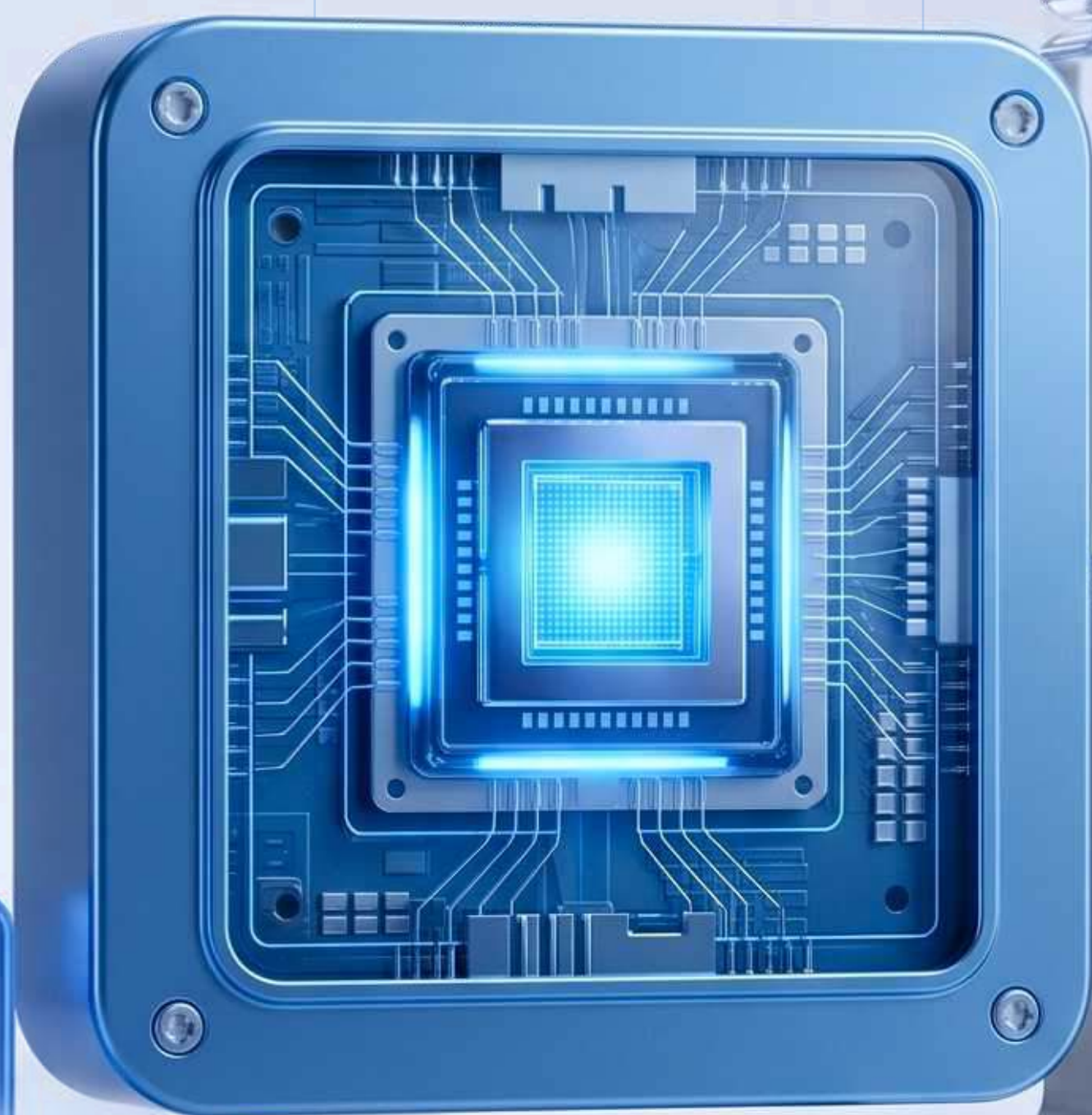
Чем больше AI-продуктов доходит до промышленного внедрения, тем выше спрос на инфраструктуру, необходимую для их развёртывания и эксплуатации. Рост спроса увеличивает загрузку инфраструктуры и выручку провайдеров, повышая экономическую обоснованность дальнейших инвестиций в её развитие.



Источники: World Bank, GDP, PitchBook, Dealroom, OECD, анализ KAMA FLOW.



Экономика AI-решений



Экономика AI-решений

Эволюция бизнес-модели. AI меняет привычную SaaS-экономику: у продукта появляется переменная себестоимость

В классической SaaS-модели основные затраты возникают на этапе разработки, после чего один и тот же продукт можно многократно продавать новым клиентам почти без увеличения себестоимости. Это обеспечивает высокую валовую маржу – 70-85% – и масштабируемость бизнеса.

В AI-native продуктах каждый запрос создаёт переменные расходы на inference, а внедрение требует адаптации под данные и процессы клиента. В результате валовая маржа снижается до 45-52%, а бизнес совмещает роли вендора и интегратора.

Критерий	Лицензия	SaaS	AI-native
Объект монетизации	Право на использование ПО	Доступ к облачному продукту	Результат / автоматизация задачи
Модель выручки	Разовая продажа, после — дополнительные продажи поддержки	Подписка за пользователя / месяц	Гибрид: подписка с оплатой по факту использования. 92% AI-компаний используют гибридную модель
Формирование себестоимости	Внедрение и поддержка у клиента	Затраты на облако, поддержка	Инференс, API / GPU. Переменная стоимость обработки каждого запроса пользователя
Валовая маржинальность	60% без учёта внедрения	70–85%	45–52%
Как масштабируется бизнес	Через увеличение числа внедрений	Через рост клиентской базы и повышение retention	Через рост клиентской базы и числа запросов, снижение стоимости их обработки
Риски	Длинный цикл продаж, сложная кастомизация и интеграция, нерегулярная выручка	Плохая юнит-экономика: высокий Churn и CAC	Высокий COGS, зависимость от моделей, сложности с расчётом окупаемости

В AI-native продуктах рост использования не всегда приводит к увеличению маржинальности. Чем активнее клиент использует продукт, тем выше могут быть расходы на inference.

Внедрение как часть бизнес-модели

Помимо переменной стоимости inference, экономика AI-продуктов осложняется высокой долей внедрения и кастомизации. Значительная часть затрат возникает при адаптации решения под данные, процессы и IT-контуры каждого клиента. Поэтому на ранних этапах AI-компания часто развивается как гибрид продукта и интегратора.

«Для AI в B2B классическая SaaS-модель работает плохо: это скорее полусервисные истории, которые всегда требуют кастомизации. В AI-решениях очень важна внедренческая часть. Когда команда сделала несколько внедрений, она начинает видеть, что из этого можно продуктивизировать».

Артур Самигуллин
Руководитель платформы Yandex AI Studio



Структура выручки влияет на мультипликатор оценки

Гибридная модель снижает потенциальную оценку компании – если значительная часть выручки приходится на внедрение и кастомную разработку, инвесторы оценивают компанию скорее как интегратора, а не как SaaS-бизнес.

«На российском рынке почти нет чисто продуктовых компаний. Обычно это гибрид: продукт, интеграция, кастомная разработка и проектная деятельность. Инвестор смотрит на бизнес, в котором много интеграционной и проектной выручки, и применяет мультипликатор интегратора. А основатели хотят, чтобы к нему применили продуктовый SaaS-мультипликатор – желательно ещё и западный. Это одна из главных проблем рынка».

Алексей Постернак
Директор по инвестициям и развитию MWS



Рынок фундаментальных моделей движется к снижению стоимости инференса при росте качества, а технологический разрыв между ведущими разработчиками постепенно сокращается

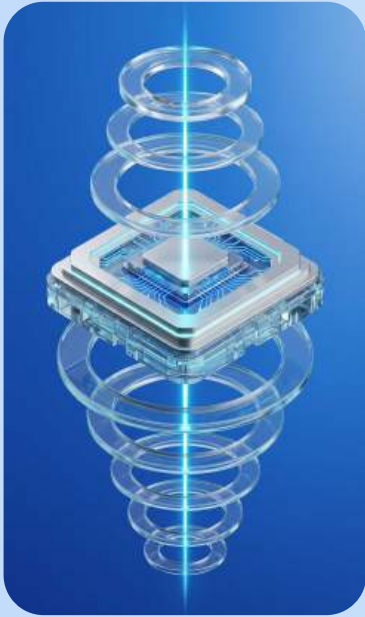
На фоне роста конкуренции и эффективности моделей ключевым критерием выбора становится не абсолютная цена, а соотношение стоимости и производительности.



На основе данных LLM Stats мы выделили ключевые технологические и экономические тенденции рынка фундаментальных моделей:

Экономика инференса улучшается

Стоимость инференса моделей уровня GPT-4 резко снизилась: в 2023 году сопоставимый уровень стоил около \$30 за 1 млн токенов, сегодня – менее \$1. LLM Stats оценивает темп снижения в 10–100 раз в год. Удешевление обеспечивается усилением конкуренции, ростом эффективности моделей и развитием вычислительной инфраструктуры.



Малые модели догоняют большие по эффективности параметров

Модель на 7 млрд параметров сегодня может показывать результаты, для которых год назад требовалась архитектура более чем на 70 млрд. За год необходимый масштаб сократился примерно на порядок при сопоставимом качестве. Это делает возможным запуск сильных моделей локально, в том числе на пользовательском оборудовании, и снижает требования к инфраструктуре при развёртывании.



Разрыв между открытыми и закрытыми моделями сокращается

Открытые модели Llama, Mistral и Qwen уже могут соответствовать GPT-4 или превосходить его на отдельных тестах. По оценке LLM Stats, open-source модели обычно отстают от закрытых frontier-моделей на 6–18 месяцев, однако этот временной разрыв сокращается.



Состав лидеров модельных рейтингов быстро обновляется

Модели, которые возглавляли бенчмарки полгода назад, сегодня могут находиться уже в середине рейтинга.



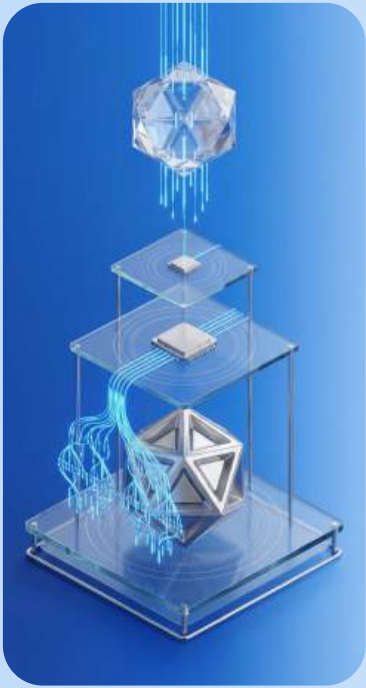
США сохраняют лидерство, Китай сокращает отставание

Американские разработчики – OpenAI, Anthropic и Google – по-прежнему лидируют в большинстве бенчмарков. При этом китайские лаборатории, включая DeepSeek, Alibaba и ByteDance, быстро сокращают разрыв, особенно в задачах логического рассуждения и программирования.



Рынок будет двигаться к многоуровневому прайсингу и маршрутизации

Использование одной модели для всех запросов становится экономически неэффективным. Большинство типовых задач выгоднее обрабатывать небольшими и дешёвыми моделями, направляя сложные запросы в более дорогие frontier-модели. Поэтому ключевым преимуществом AI-продукта становится способность маршрутизировать запросы между моделями, балансируя качеством, скоростью и стоимостью инференса.



Экономика frontier-моделей пока зависит от существенного дотирования вычислительных мощностей, а пользовательская цена AI-продуктов не отражает их фактическую стоимость

«Подписка Anthropic стоимостью \$200 позволяет пользователю потребить вычислительных ресурсов примерно на \$26 тыс. Это показывает масштаб дотирования рынка».

Армен Амирханян
Директор по развитию искусственного интеллекта Московской биржи



Развитие базовых моделей способно за считанные месяцы радикально изменить экономику AI-решений. Например, в одном из кейсов себестоимость одного и того же процесса снизилась на два порядка всего за полгода.

«В одном из проектов ручная обработка документа обходилась примерно в 50 рублей, использование GPT-4 снизило стоимость до 30 рублей, а гибридная схема с GPT-3.5 — до 15 рублей. Через полгода новое поколение моделей позволило сократить себестоимость примерно до 50 копеек».

Евгений Кокуйкин
CEO стартапа HiveTrace



Снижение удельной стоимости инференса не гарантирует снижения общих расходов

Парадокс Джевонса: чем дешевле использование ресурса, тем быстрее растёт его потребление.

Агентные системы разгоняют потребление токенов

Особенно быстро потребление растёт в агентных системах: вместо одного ответа модель выполняет цепочку действий — планирует, обращается к инструментам, анализирует результаты, проверяет себя и повторяет цикл.

Многие этапы требуют отдельного вызова модели и повторной обработки накопленного контекста. Поэтому даже при снижении стоимости одного вызова общее потребление токенов и спрос на вычислительные мощности могут расти.

5–30x

По оценке Gartner, выполнение одной задачи агентной моделью требует в 5–30 раз больше токенов, чем взаимодействие со стандартным чат-ботом.

В результате объём потребления токенов и спрос на вычисления могут расти быстрее, чем снижается их удельная стоимость.



Планирование



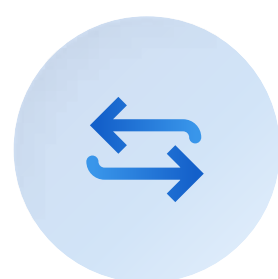
Инструменты



Анализ результатов



Проверка результата



Повтор цикла

Средний запрос становится значительно тяжелее

По данным трафика платформы OpenRouter (агрегатор и единый API-шлюз к LLM) с 2024 года:

4x

Средний объём входного контекста вырос примерно в 4 раза – с 1,5 тыс. до более чем 6 тыс. токенов.



3x

Средний объём ответа вырос почти в 3 раза – примерно со 150 до 400 токенов.



>3x

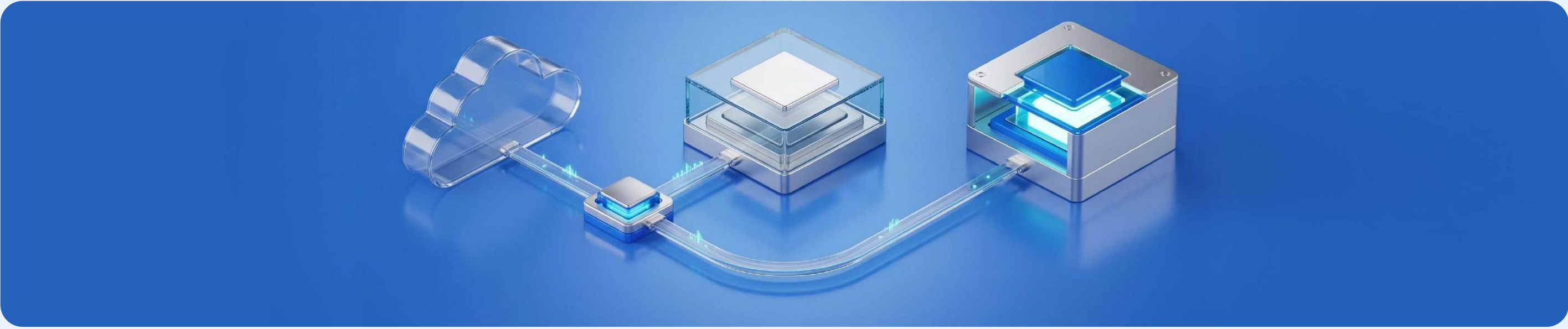
Средняя длина последовательности, отражающая сложность задачи и глубину взаимодействия, выросла более чем в 3 раза за 20 месяцев – с менее 2 тыс. до более 5,4 тыс. токенов.



50%+

К концу 2025 года reasoning-модели обеспечивали уже более половины всего токенового трафика платформы.





Выбор архитектуры развертывания определяется тремя ключевыми факторами: масштабом нагрузки, регуляторными требованиями и наличием ML-команды

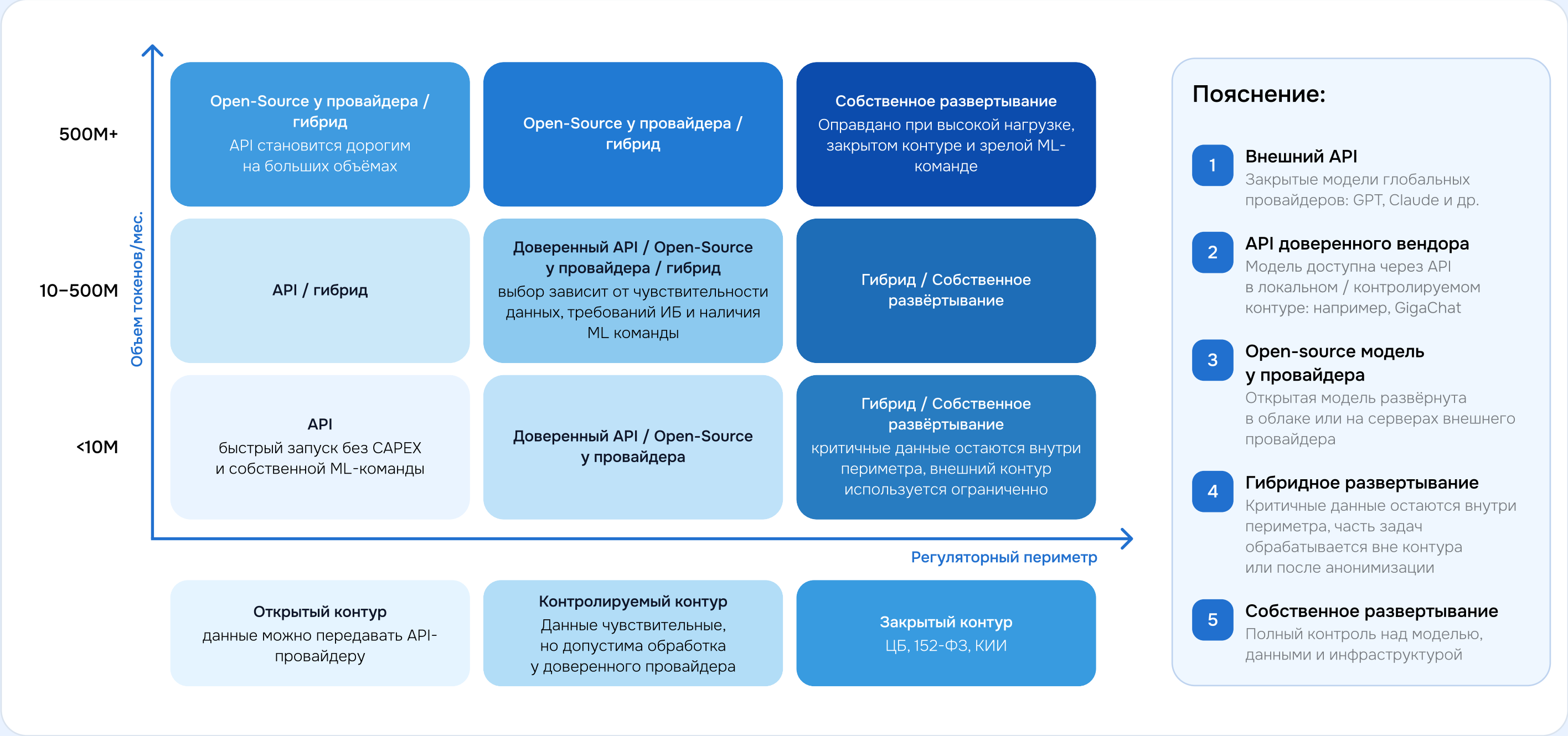
При внедрении AI-решений компании сталкиваются с базовым архитектурным выбором – использовать внешний API, развернуть open-source-модель у провайдера или построить собственный контур.

В России особенно распространено развёртывание моделей на собственной инфраструктуре – прежде всего из-за требований к безопасности данных, регуляторных ограничений и стремления сохранить контроль над технологическим контуром. Однако выбор архитектуры также должен зависеть от масштаба нагрузки и наличия компетентной ML-команды.

«Облако позволяет быстро запустить пилот и проверить гипотезу, но по мере роста потребления операционные расходы увеличиваются.

После подтверждения эффекта на этапе масштабирования решение можно переносить в локальную инфраструктуру: капитальные затраты становятся выше, зато стоимость инференса фактически сводится к стоимости электроэнергии».

Алексей Бобок
СРО AI4BI



Команда – главный скрытый драйвер издержек внедрения

Собственный контур снижает зависимость от внешнего провайдера, но переносит инфраструктурную и операционную сложность внутрь компании. Модели требуют постоянного обновления, мониторинга инфраструктуры обработки запросов и безопасности.

Всё это требует труда квалифицированных ML-инженеров. Так, медианная зарплата ML-разработчика в России составляет около 230 тыс. рублей в месяц на руки, а специалистов senior-уровня – порядка 400 тыс. рублей.

В итоге стоимость собственного развёртывания измеряется не столько затратами на «железо», сколько постоянным ФОТ специалистов.

Источники:
Хабр Карьера; материалы RAFT.

Собственное развёртывание

3–6+ инженеров: ML/MLOps, DevOps, GPU-инфраструктура, безопасность

Развёртывание open-source на инфраструктуре провайдера

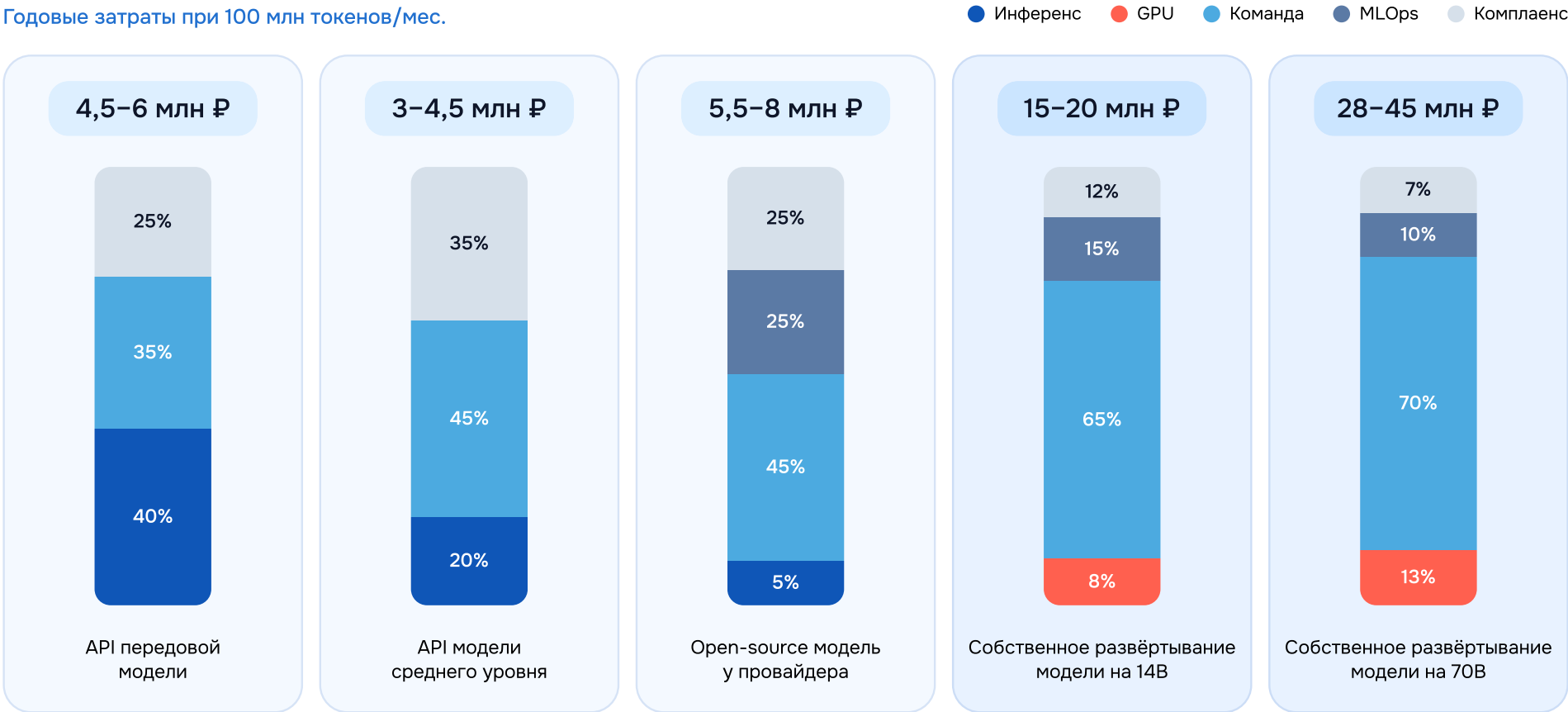
2–3+ инженера: ML, интеграция, мониторинг, обновления

API

1–2 инженера на интеграцию и контроль качества

Собственное развёртывание моделей окупается только при высокой нагрузке, полноценной ML-команде и жёстких требованиях к безопасности

Структура затрат годового обслуживания при разных сценариях архитектуры развёртывания:



«Основные затраты приходятся на вычислительные мощности и весь жизненный цикл оборудования — от поставки до сопровождения.

Команда также остается крупной статьёй расходов, потому что устоявшейся методологии еще нет и решения требуют значительной ручной доработки».

Антон Шмаков
Технический директор,
ПАО Группа Астра



«В крупных корпоративных проектах до 80% затрат может приходиться не на оборудование, а на создание хранилища, интеграцию, настройку пайплайнов, нормализацию данных и формирование управляемого аналитического слоя. Само оборудование в такой структуре может составлять около 20% стоимости проекта».

Алексей Бобок
СРО AI4BI



API

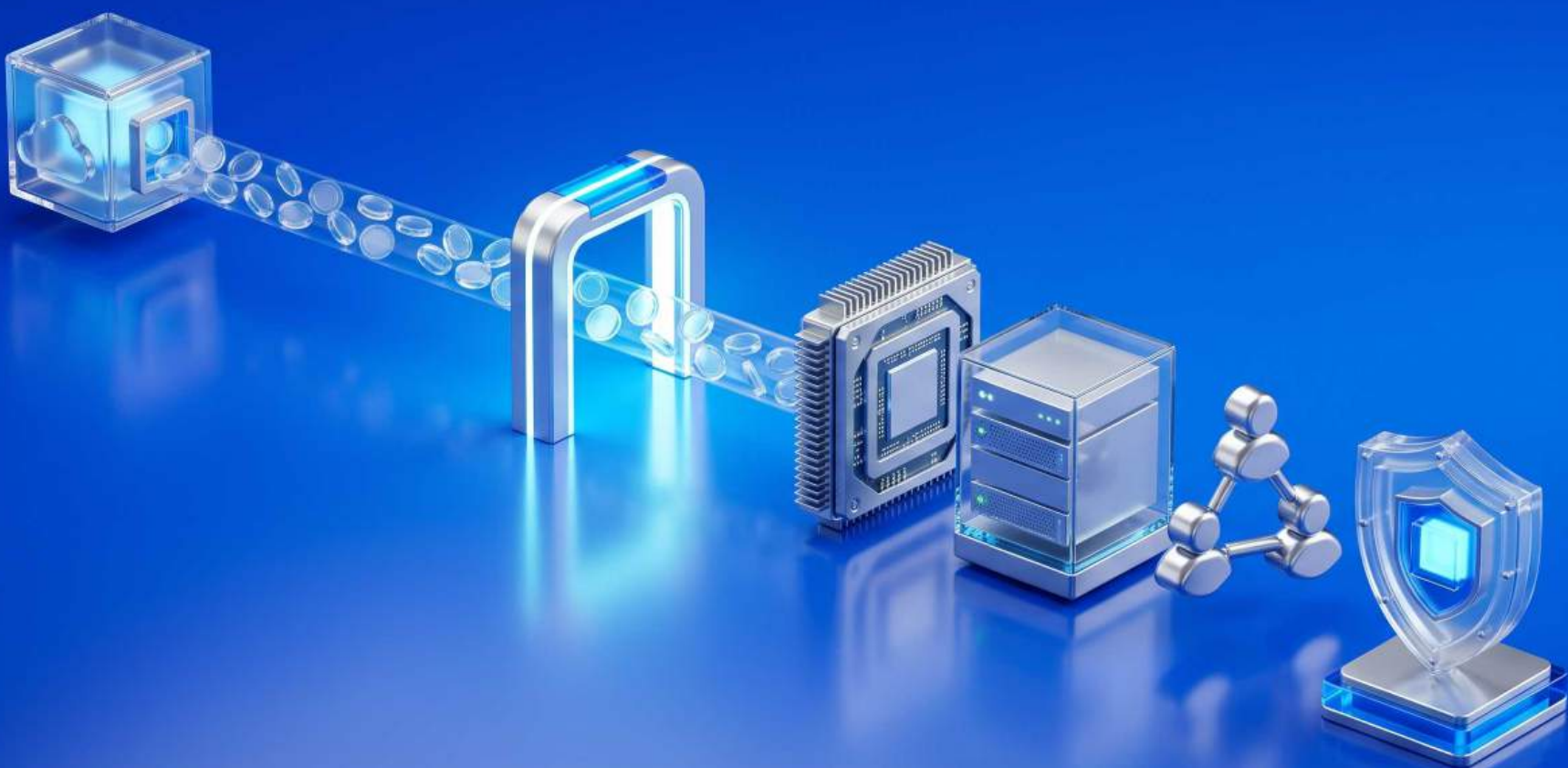
Быстрый запуск и низкая операционная сложность, но зависимость от стоимости токенов и провайдера.

Open-source модель у провайдера

Ниже расходы на инференс, но появляются затраты на MLOps, мониторинг, обновления и контроль качества.

Собственное развёртывание

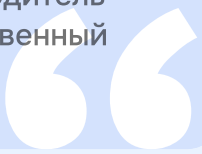
Максимальный контроль над данными и инфраструктурой, но высокие фиксированные расходы на команду, GPU, эксплуатацию и безопасность.



При переходе от API к собственному развёртыванию меняется структура затрат – вместо оплаты токенов растут расходы на команду, GPU-инфраструктуру, MLOps и обеспечение безопасности

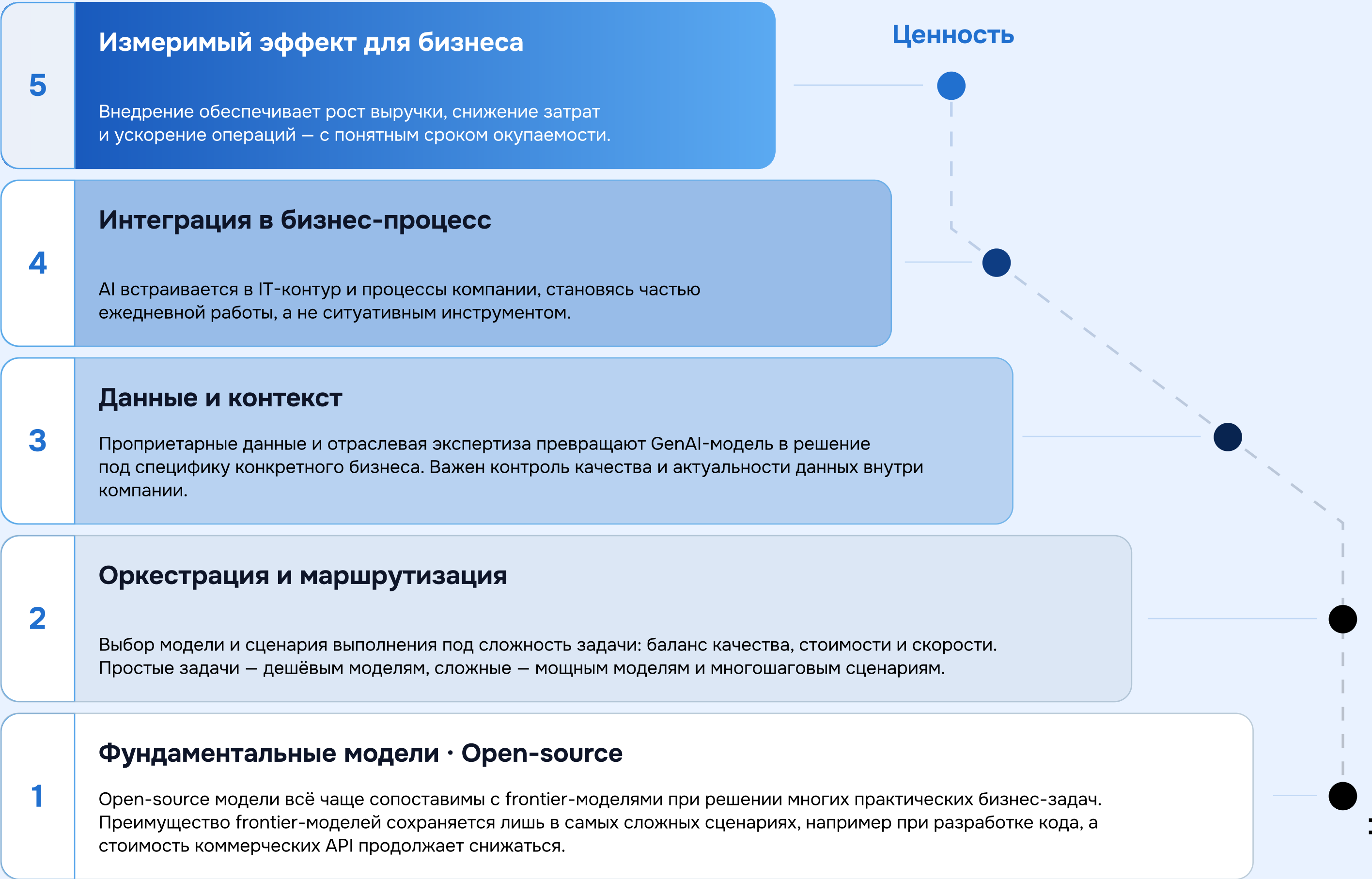
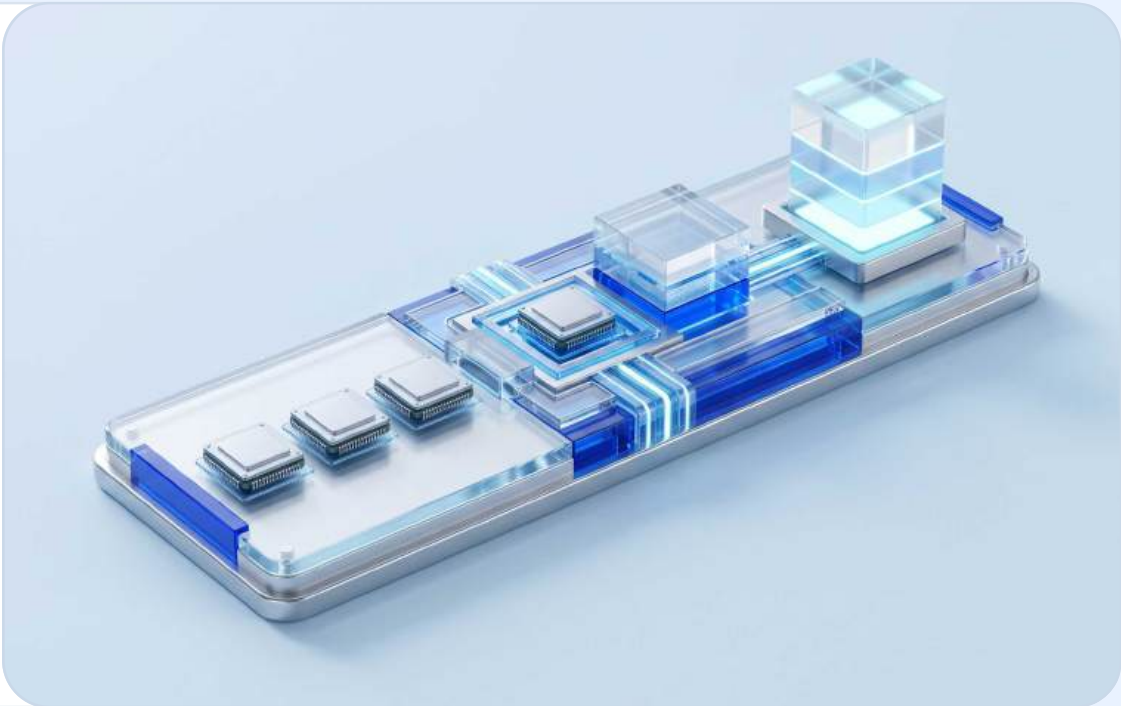
«Мы часто сталкиваемся с ситуацией, когда компания готова потратить на пилот небольшую сумму – 2–3 млн рублей: кейс понятен, ожидаемый эффект тоже. Но инфраструктура под такой пилот может стоить уже около 8 млн рублей – экономика пилота ломается ещё до проверки гипотезы. Это становится несоразмерно риску, когда компания ещё не до конца понимает, зачем ей это нужно и даст ли внедрение результат».

Лариса Малькова
Управляющий директор, Руководитель практики «Прикладной Искусственный Интеллект», Axenix



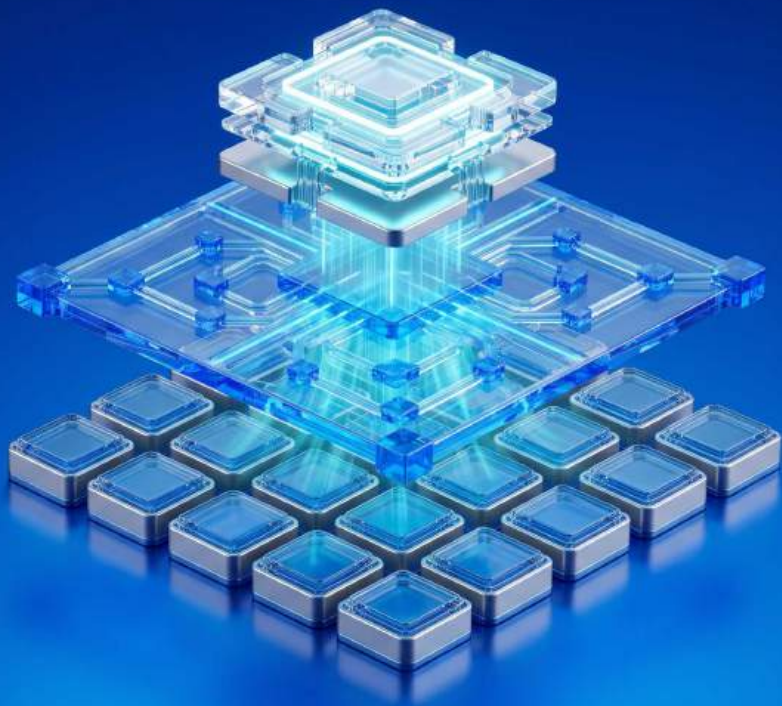
Фундаментальные модели становятся взаимозаменяемыми в ряде задач. Ценность решения создаётся в данных, интеграции в процесс и измеримом бизнес-эффекте

Интерес представляют не «обёртки над LLM», а продукты, где есть подтверждённая экономика внедрения и масштабируемость от клиента к клиенту.



«Максимальный эффект возникает на пересечении трёх факторов: наличия массивов данных, цифровизированного процесса и высокой ценности самой функции для бизнеса».

Антон Шмаков
Технический директор,
ПАО Группа Астра



По мере того как базовые модели становятся доступнее и меньше отличаются друг от друга, основная прибыль смещается в прикладные решения и интеграцию

«Базовые LLM — это инфраструктурный фундамент для создания конечных AI-продуктов. Экономика провайдеров таких моделей будет строиться прежде всего на масштабе и частотности их использования, тогда как основная маржа будет формироваться в прикладном слое — у отраслевых разработчиков и интеграторов, адаптирующих решения под конкретные бизнес-процессы, данные и специфику клиентов».

Александр Вибе
Директор Центра
ИИ-решений для B2B
СберБанка



Цепочка создания ценности AI-решений: мнения участников рынка

Измеримый эффект для бизнеса

Продавать не AI, а экономический результат

«Наиболее сильную позицию займёт тот, кто продаёт не просто AI-решение, а измеримое улучшение бизнеса — например, снижение затрат на 10%. В таком случае компания фактически продаёт операционную эффективность с понятным для клиента сроком окупаемости».

Виктор Поздняков
Руководитель направления стратегических инвестиций Т-Банка

Компании не умеют считать окупаемость AI-решений

«90% инициатив заканчиваются на столе у финансового директора. Не потому, что тот плохой, а потому, что он просто не знает, как считать такие проекты».

Александр Ведяхин
Первый заместитель председателя
правления Сбербанка
Интервью Forbes, 24 февраля 2026 года

Заявленный эффект необходимо подтверждать метриками

«Компании заявляют о значительных эффектах от AI, но пока эти деньги сложно увидеть. Складывается впечатление, что у всех есть KPI на внедрение, поэтому компании показывают отдельные кейсы и заявляют об огромной экономии. Однако проверить, сколько бизнес реально сэкономил или заработал благодаря AI-агенту, зачастую невозможно: нормальных метрик для измерения этих эффектов пока нет».

Алексей Постернак
Директор по инвестициям и развитию MWS

Рост производительности от AI не всегда напрямую отражается в P&L — эффект необходимо считать через дополнительный объём работы и предотвращённый рост штата

«Если разработчик стал работать в два раза быстрее, расходы компании не сокращаются автоматически. Но команда может выполнять больше задач без пропорционального расширения штата. Поэтому эффект AI важно считать по косвенным факторам: сколько дополнительных специалистов пришлось бы привлечь, чтобы выполнить тот же объём работы».

Армен Амирханян
Директор по развитию искусственного
интеллекта Московской биржи

«Все чувствуют, что AI повышает эффективность, но бизнесу эту эффективность сложно измерить. Сотрудники могут быстрее выполнять задачи, но в финансовых результатах компании это необязательно отражается».

Армен Амирханян
Директор по развитию искусственного
интеллекта Московской биржи

Интеграция в бизнес-процесс

Бизнесу нужен не ещё один IT-продукт, а AI внутри существующих систем

«Для стартапов лучшая стратегия — встраиваться в экосистемы крупных вендоров. Промышленники не хотят покупать «ещё одну IT-систему». Более жизнеспособный путь — AI-модуль внутри ERP, ЭДО, SCADA, 1C или платформы крупного интегратора».

Андрей Фиохин
Генеральный директор ООО «Айтисфера» (входит в ГК «Азот»)

От автоматизации отдельных операций — к пересборке бизнес-процесса целиком

«Реальный эффект от AI появляется, когда компания перестраивает бизнес-процесс end-to-end и внедряет технологию в операционную логику».

Артур Самигуллин
Руководитель платформы Yandex AI Studio

Компании внедряют не AI как таковой, а инструменты повышения эффективности конкретного бизнес-процесса

«В IBS на текущий момент мы рассматриваем искусственный интеллект прежде всего как инструмент повышения эффективности, встраиваемый в процессы, а не как stand-alone инвестиционный объект: мы внедряем AI-ассистентов и агентные решения во внутренние процессы разработки и тестирования, а также реализуем прикладные кастомные проекты для заказчиков. Наш опыт показывает, что ценность возникает там, где технология встроена в конкретный бизнес-процесс, опирается на качественные данные и обеспечивает измеримый экономический эффект при требуемом уровне надежности и информационной безопасности».

Анна Буракова
Руководитель M&A IBS

Ценность AI-решения создаёт интеграция с корпоративным контуром

«Рынок требует готовых индустриальных агентов с глубокими знаниями для конечных пользователей. Сама по себе LLM ничего не знает про тебя, твою компанию, процессы и регламенты. Если хочешь агентскую историю, она должна работать с корпоративными системами».

Лариса Малькова
Управляющий директор, Руководитель практики
«Прикладной Искусственный Интеллект», Axenix

Цепочка создания ценности AI-решений: мнения участников рынка

Данные и контекст

Доступ к данным становится главным конкурентным преимуществом

«Гонка уже не столько в том, у кого лучше модель или алгоритм, сколько в том, какие данные использовались для обучения. Есть очевидные отрасли, где данные доступны, например LegalTech.

Но существует множество закрытых и менее очевидных данных внутри промышленных и традиционных компаний. Именно там может находиться значительный потенциал для создания новых продуктов».

Алексей Постернак
Директор по инвестициям и развитию MWS

Отраслевые решения требуют партнёрства с владельцами данных

«Для создания отраслевого решения необходимо договориться с компанией о предоставлении данных.

Скорее всего, такие компании захотят участвовать в проекте не только как заказчики, но и как совладельцы».

Алексей Постернак
Директор по инвестициям и развитию MWS

Стоимость AI-проекта определяется зрелостью данных и процессов работы с ними

«Если данные структурированы, актуальны и регулярно обновляются, стоимость самого AI-проекта может быть относительно небольшой.

На практике выясняется, что данных не хватает, они не структурированы, а процессов работы с ними нет – поэтому дорогим оказывается весь проект».

Антон Шмаков
Технический директор, ПАО Группа Астра

Оркестрация и маршрутизация

Развертывание open-source моделей требует полноценной инфраструктуры для эксплуатации и формирует спрос на решения для оркестрации, мониторинга и оптимизации вычислительных ресурсов.

«Эксперименты с open-source всё равно требуют продуктивизации – управления рисками, жизненным циклом и развитием решения. Поэтому для инфраструктурных компаний останется большое поле: оптимизация ресурсов, распределение нагрузки, мониторинг и управление вычислительными узлами».

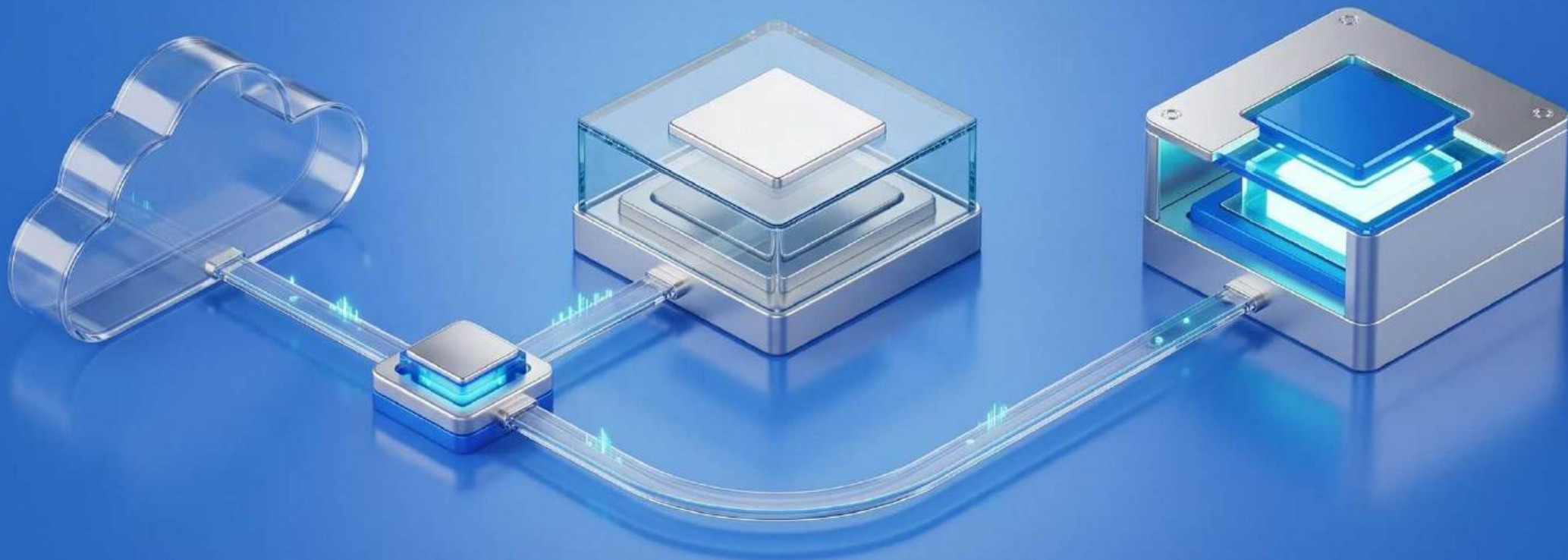
Антон Шмаков
Технический директор ПАО «Группа Астра»

Фундаментальные модели

Модель — лишь технологическая основа, а не самостоятельный продукт

«Ключевые конкурентные преимущества России лежат в создании готовых продуктов, а не базовых технологий. Необходима не просто технология — технологию нужно обернуть в продукт».

Андрей Рыбинцев
Управляющий директор по ИИ Авито



Российские AI-стартапы:

ограничения
и условия роста



Российские AI-стартапы: ограничения и условия роста

Build vs Buy: российские корпорации чаще выбирают внутреннюю разработку, чем решения AI-стартапов

M&A-кандидат должен обладать активом, который невозможно быстро воспроизвести внутри корпорации: технологией, данными, отраслевой экспертизой или глубоким пониманием клиента.

«Если раньше частью инвестиционного тезиса для покупки компании могла быть уникальная технология или продукт, то теперь мы каждый раз задаёмся вопросом: действительно ли компании необходимо покупать или аналогичное решение можно быстро разработать внутри?»

Виктор Поздняков
Руководитель направления
стратегических инвестиций Т-Банка

“

Почему компании предпочитают разрабатывать AI-решения самостоятельно?

Генеративный AI ускоряет внутреннюю разработку

С помощью AI-инструментов корпоративные команды могут быстро собрать MVP (минимально жизнеспособный продукт) для проверки гипотез без покупки готового решения.

«Барьеры на разработку уменьшаются и многие решения можно быстро прототипировать, но критично умение продавать, обслуживать и удерживать клиентов — то, что нельзя «завайбкодить» за неделю».

Виктор Поздняков
Руководитель направления
стратегических инвестиций Т-Банка

“

AI-кодинг уже переходит от пилотов к массовому использованию внутри крупнейших российских корпораций

«Корпоративный рынок AI-кодинга в России перспективен: уже больше половины кода внутри Сбера пишется с использованием GigaCode».

Александр Вибе
Директор Центра ИИ-решений для B2B
СберБанка

“

Доступ к вычислениям концентрируется у крупных игроков

Для сложных AI-сценариев критично наличие вычислительных мощностей.

≈80%

поставок AI-ускорителей в России приходится на крупнейших экосистемных игроков
Оценка T1, 2025 год

У крупных компаний уже есть база для внедрения AI

Накопленные данные, внутренняя ML-команда и понимание бизнес-процессов позволяют быстрее адаптировать решение под собственный контур.

Рынок внешних AI-стартапов остается незрелым

Крупные компании отмечают низкую зрелость команд, легкость воспроизводимости продуктов и недоказанный экономический эффект внедрения.

«С точки зрения M&A нам неинтересна покупка небольших универсальных «AI-оболочек» над общедоступными моделями: такие продукты относительно легко воспроизводимы и могут быстро превратиться в стандартную функцию крупной платформы. При этом мы готовы рассматривать компании, обладающие собственной технологической экспертизой, доступом к уникальным данным, подтвержденными внедрениями и возможностью тиражировать решение в рамках отрасли».

Анна Буракова
Руководитель M&A IBS

“

«Мы близки к тому, чтобы самостоятельно закрывать весь цикл необходимых AI-продуктов, но продолжаем следить за рынком и готовы покупать сильные локальные решения. Пока многие задачи мы можем качественно решить внутри, иностранные продукты недоступны, а российские по ряду направлений ещё не достигли необходимой зрелости».

Алексей Постернак
Директор по инвестициям и развитию MWS

“

AI-стартапы конкурируют не только между собой, но и с внутренней разработкой заказчика



AI-стартапы в России упираются в дефицит ключевых ресурсов: вычислительные мощности, капитал и данные

Стартапы в Enterprise Software традиционно выигрывали у крупных компаний за счёт скорости разработки и гибкости. Однако с развитием AI этого уже недостаточно – генеративные инструменты резко снизили барьер для создания MVP, поэтому способность быстро собрать решение перестаёт быть устойчивым конкурентным преимуществом.

Для захвата рынка стартапу нужны более сильные основания – прежде всего доступ к уникальным отраслевым данным, глубокое понимание бизнес-процессов и вычислительные ресурсы для развёртывания и масштабирования продукта.

Одним из наиболее труднодоступных ресурсов для независимого стартапа становятся данные. Так, стартап может использовать открытые источники или приобрести готовый датасет. За рубежом доступ к внешним данным упрощают специализированные маркетплейсы. Например, AWS Data Exchange объединяет более 3,5 тыс. дата-продуктов от свыше 300 поставщиков.

В России необходимость формирования подобной инфраструктуры признаётся ключевыми участниками рынка, однако масштабный механизм коммерческого обмена данными пока не сложился. Проблема неоднократно обсуждалась на отраслевых площадках: в частности, на Data Fusion в 2025 году обсуждались идеи создания национальной биржи данных и зарубежные модели их коммерциализации.

Тем не менее доступ стартапов к качественным отраслевым датасетам преимущественно зависит от прямых договорённостей с заказчиками. Получить данные стартап может в рамках пилота с корпоративным заказчиком или постепенно накопить их в процессе эксплуатации собственного продукта. Однако компании неохотно выводят чувствительную информацию за пределы собственного контура из-за требований информационной безопасности и рисков утечек, необходимости защиты персональных и коммерческих данных, а также неясности прав на дальнейшее использование датасетов и обученных на них моделей.

При этом, как было показано ранее, корпорации в целом осторожно относятся к решениям независимых стартапов — заказчики не всегда видят измеримый бизнес-эффект, сомневаются в зрелости продукта и зачастую предпочитают внутреннюю разработку. В результате формируется замкнутый круг — без корпоративного заказчика стартапу сложно получить данные и доработать решение, а без готового зрелого решения с подтвержденным бизнес-эффектом — привлечь корпоративного заказчика.

В результате российские AI-стартапы оказываются в ситуации дефицита трёх взаимосвязанных ресурсов: вычислительных мощностей, капитала и спроса, а также доступа к данным.



«Мы очень инженерная страна: в России почти в каждой компании есть ощущение, что AI-решение можно сделать внутри».

Артур Самигуллин
Руководитель платформы Yandex AI Studio

Корпоративный заказчик одновременно выступает для стартапа и потенциальным партнёром, и конкурентом.

Корпорации располагают собственными командами разработки, накопленными данными, вычислительными кластерами и доступом к внутренним бизнес-процессам, что позволяет им самостоятельно развёртывать AI-решения.

Одновременно небольшой объём российского венчурного рынка ограничивает возможности привлечения капитала для развития.

Финансирование необходимо в том числе для аренды вычислительных мощностей, а при создании собственной инфраструктуры — для значительных единовременных вложений в оборудование.



«Все AI-стартапы хотят получить токены или GPU — многие готовы брать вычислительные мощности вместо денег».

Игорь Маслов
Главный технический директор Т-Банка

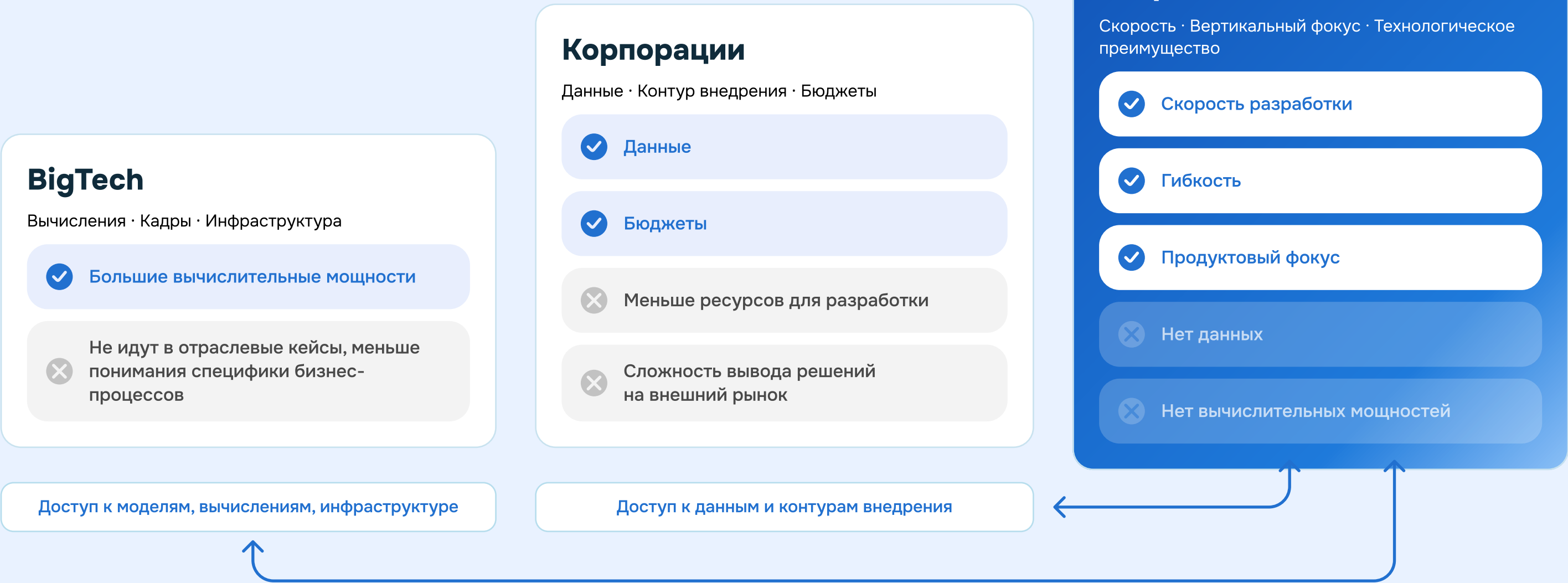
Модель взаимодействия для формирования рынка независимых AI-решений

Для развития рынка независимых AI-решений мы видим следующую модель взаимодействия между участниками рынка. Под BigTech в данном случае понимаются крупнейшие IT-компании, которые развивают фундаментальные модели и облачную инфраструктуру.

Корпорации – это крупные отраслевые заказчики, например, промышленные нефтегазовые компании, ритейл и т. д., внутри которых сосредоточены данные, подходящие для автоматизации бизнес-процессов, и потенциальный платежеспособный спрос на AI-решения.

Стартапу не обязательно самостоятельно владеть всем технологическим стеком. BigTech может предоставить ему доступ к моделям, вычислительным мощностям и базовой инфраструктуре, а отраслевой заказчик – к данным и контуру для внедрения, а также бюджетам на развитие продукта.

Стартап выступает связующим звеном, превращая эти ресурсы в прикладной продукт за счёт высокой скорости разработки, гибкости, чёткого продуктового фокуса и глубокой специализации на конкретном отраслевом сценарии.



Инвестиционная возможность

Прикладные B2B-решения в нишах, где BigTech не фокусируется, а корпорациям не хватает скорости для собственной разработки.

Инфраструктура BigTech становится основой рынка, но конечная ценность создаётся в отраслевых продуктах – через партнёрства и M&A

«Одной инфраструктуры недостаточно: корпоративному клиенту нужен конкретный отраслевой value proposition – решение, которое позволит повысить эффективность производства, добычи или переработки. Поэтому логичным направлением для бигтехов станет вертикальная интеграция через партнёрства и M&A с компаниями, обладающими глубокой отраслевой экспертизой».

Александр Вибе
Директор Центра ИИ-решений для B2B СберБанка

Будущее независимых российских AI-игроков – в отраслевых решениях, адаптированных под отечественную инфраструктуру и регулирование

«Наиболее перспективная стратегия для российских небольших AI-компаний, на наш взгляд, заключается не в конкуренции с глобальными разработчиками фундаментальных моделей или с российскими технологическими лидерами, а в создании суверенных отраслевых решений – локализации и доработке технологий с учетом возможностей AI, российской инфраструктуры, требований регуляторов и практических задач заказчиков на достаточно ёмких рынках».

Анна Буракова
Руководитель M&A IBS

Карта инвестиционных возможностей



Карта инвестиционных возможностей

Структура AI-рынка: от капиталоемких железа и базовой инфраструктуры к прикладным решениям

Российский AI-рынок характеризуется концентрацией ресурсов в капиталоемких сегментах – вычислительной инфраструктуре и фундаментальных моделях, развитие которых сосредоточено внутри IT-корпораций.

Чтобы определить, где в сложившейся конфигурации остаётся пространство для независимых компаний и частного капитала, рассмотрим цепочку создания стоимости на AI-рынке и его структуру.

Чем выше слой, тем выше барьер входа по капиталу и тем сильнее концентрация у крупнейших игроков. Верхние слои – «чистое железо», облака и модели – требуют максимального CAPEX, длинного горизонта инвестиций и фактически закрыты для частного инвестора без участия государства или крупных корпораций.

Именно поэтому фокус нашего исследования смещается вниз по стеку – к прикладным B2B-решениям.

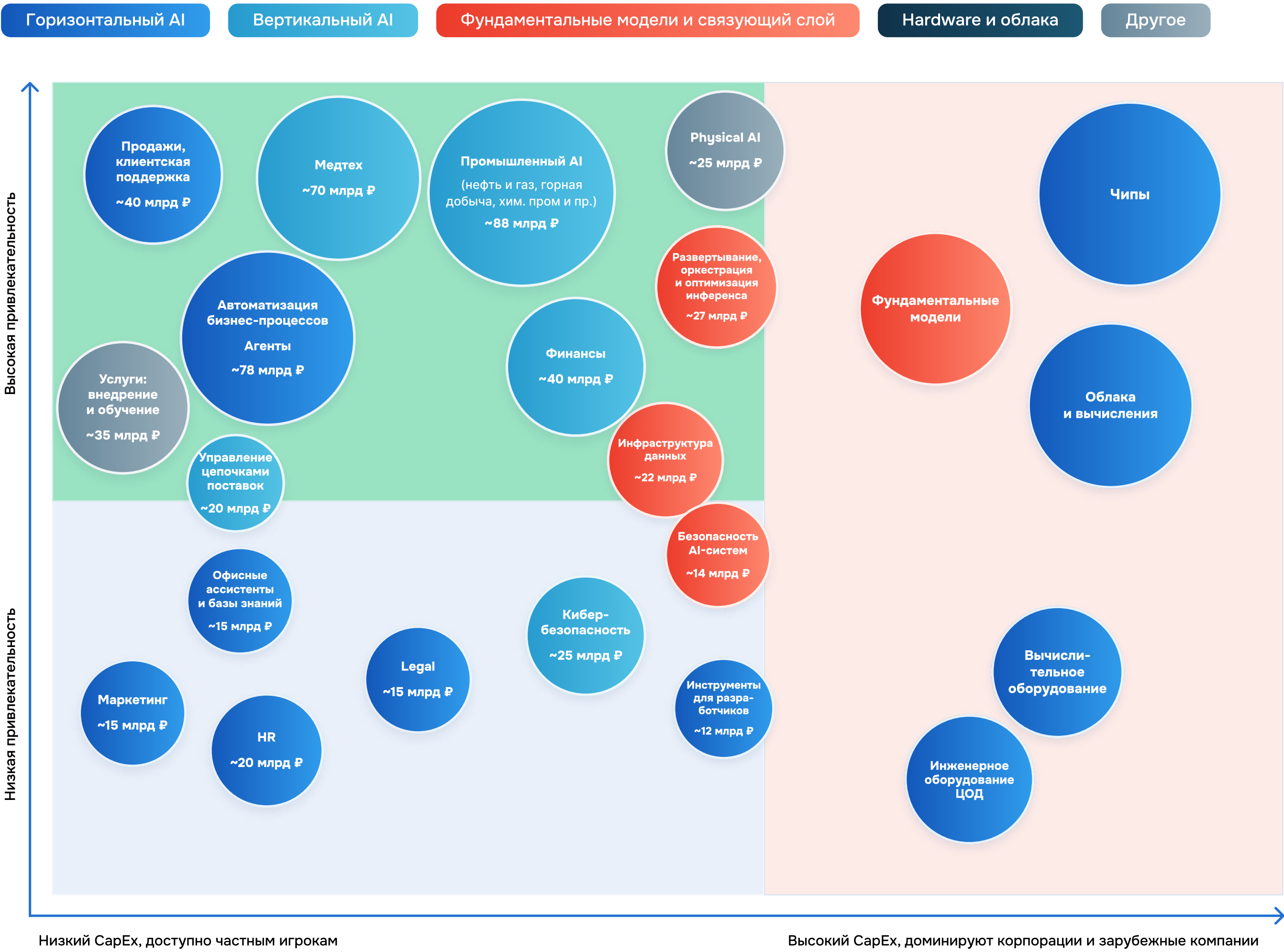
01 Hardware	Чипы • процессоры и AI-ускорители: CPU, GPU, TPU, NPU • память: HBM, DRAM, NAND • сетевые чипы Сегмент полностью занят мировыми лидерами – бенефициарами AI-бума (NVIDIA, AMD и др.). Россия критически отстаёт в передовых полупроводниках – нет места для нового локального производителя.	Вычислительное оборудование • Серверы • Сетевое оборудование • Хранение данных В России уже сформирован заметный пул локальных производителей – YADRO, Аэродиск и др.	Инженерное оборудование ЦОД • Системы электроснабжения • Системы охлаждения • Серверные стойки • Системы распределения и управления питанием Рост поддерживается строительством новых ЦОД; основные возможности – в охлаждении и энергоэффективности.
02 Облака и вычисления	Гиперскейлеры и облачные провайдеры Рынок облачной инфраструктуры уже занят крупными устойчивыми игроками (Cloud.ru, «Ростелеком-ЦОД», Selectel, MWS и др.). При этом рост AI-нагрузок требует постоянного ввода новых вычислительных мощностей и расширения дата-центров.	Неооблака — специализированные GPU-облака	Колокация
03 Фундаментальные модели	Frontier OpenAI · Anthropic · Google · xAI Open-weight DeepSeek · Meta · Alibaba · Mistral AI Высокие затраты на обучение моделей; в России развитие сосредоточено исключительно внутри корпораций.	Узкоспециализированные ElevenLabs (голос) · Runway (видео) · Suno (музыка)	Российские разработчики Сбер · Яндекс · Т-технологии
04 Данные и связующее ПО Технологии для разработки, адаптации, развертывания и промышленной эксплуатации AI-моделей	Инфраструктура данных Датасеты и маркетплейсы, сбор, очистка, разметка и синтетические данные Хранение и RAG Векторные БД, embeddings, базы знаний и семантический поиск	Оркестрация Цепочки и агенты, fine-tuning, маршрутизация моделей, API и интеграция с корпоративными системами Контроль и безопасность Мониторинг, оценка качества, защита данных и управление доступом	Эксплуатация Model serving, масштабирование, оптимизация вычислений и управление нагрузкой Обилие open-source решений ограничивает монетизацию отдельных инструментов; рынок формируется вокруг комплексных платформ.
05 Прикладные решения Конечные решения для пользователей и бизнеса	Вертикальные решения Индустриальные решения Медицина · Финансы · Кибербезопасность · Нефть и газ · Металлургия и горная добыча · Ритейл и цепочки поставок · Строительство · AgroTech · Химическая промышленность Дифференциация через уникальные данные и доменную экспертизу	Горизонтальные решения Кросс-отраслевые решения для типовых бизнес-функций Продажи и поддержка клиентов · HR · Legal · Маркетинг и дизайн · Офисные ассистенты и базы знаний · Автоматизация бизнес-процессов и AI-агенты · Инструменты для разработчиков Потенциально большой объём рынка, но решения быстро становятся взаимозаменяемыми	Physical AI AI, который воспринимает окружающую среду и взаимодействует с ней через физические устройства Промышленные роботы · Беспилотный транспорт · дроны · Автономная техника · Умные устройства Перспективный передовой сегмент на ранней стадии формирования
06 Услуги Интеграция AI в бизнес-процессы и адаптация сотрудников	Внедрение и интеграция Подключение к внутренним системам и адаптация под заказчика	Обучение и консалтинг Формирование стратегии внедрения AI · Обучение сотрудников · Взаимодействие с AI-инструментами	Массовое внедрение AI потребует не только интеграции новых решений в существующий IT-стек, но и перестройки бизнес-процессов, а также адаптации сотрудников к новым инструментам и принципам работы. Это формирует значительный рынок внедрения, консалтинга и корпоративного обучения.

Карта инвестиционных возможностей на российском рынке AI-решений

Совокупный потенциал сегментов российского AI-рынка, доступных частным независимым игрокам, к 2030 году оценивается в более чем

561 млрд ₽

Команда KAMA FLOW провела анализ рынка и сопоставила ключевые рыночные сегменты по инвестиционной привлекательности, масштабу и капиталоёмкости.



Выводы для частного инвестора

Сегменты «чистого железа» и базовой AI-инфраструктуры – фундаментальные модели, облака – требуют высокого CapEx, длинного инвестиционного горизонта и часто недоступны частному инвестору без участия государства или крупных корпораций.

Инвестиционно привлекательными выглядят сегменты в «зеленой зоне» – высокий/средний барьер входа и низкий CapEx, измеримый эффект:

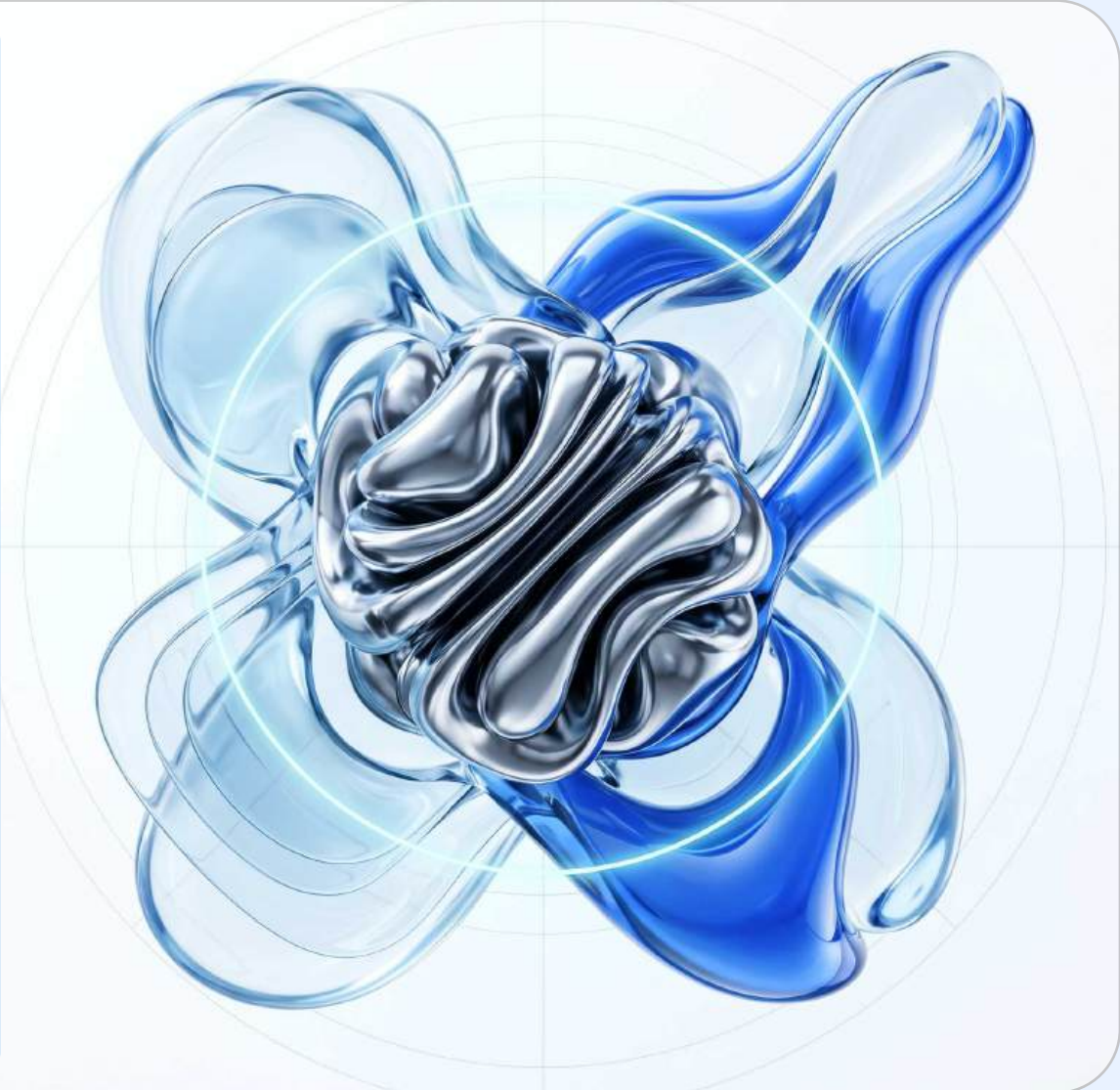
Вертикальные решения в сфере медицины, промышленности и финансов

Услуги

Physical AI

Решения в сфере данных и связующего ПО

Горизонтальные решения в сфере автоматизации продаж и агенты



Ключевые инвестиционно привлекательные сегменты российского рынка

Вертикальный AI

«Основной потенциал сейчас находится в отраслевых решениях. Крупные технологические компании не могут одновременно быть экспертами в нефтянке, металлургии, ритейле, банкинге и всех остальных индустриях. При этом небольшая команда может взять доступную open-source-модель, получить данные изнутри отрасли и построить на них новый продукт. Именно в такие истории, на мой взгляд, интересно инвестировать».

Алексей Постернак
Директор по инвестициям и развитию MWS



Промышленный AI

Отраслевые решения для нефтегаза, горной добычи, металлургии, химической промышленности и др.

88 млрд ₽

Инвестиционный тезис:

Россия – крупная индустриальная экономика с развитым нефтегазовым, горнодобывающим, металлургическим и химическим секторами. Барьер входа формирует доступ к отраслевым данным. Высокая стоимость решений компенсируется измеримым экономическим эффектом – снижением простоев, брака, аварийности и расхода сырья.

Экспертные интервью с участниками рынка подтверждают гипотезу о привлекательности рынка:

«В промышленности AI не масштабируется «из коробки»: ценность упирается в интеграцию и кастомизацию. Даже похожие производства отличаются данными, оборудованием и IT-ландшафтом: у предприятий разные ERP, ЭДО, АСУ ТП, справочники, оборудование, процессы».

Андрей Фиошин
Генеральный директор ООО «Айтисфера»
(входит в ГК «АЗОТ»)



«Более вероятный сценарий – взять сложную инженерную компанию, которая знает процессы в металлургии или промышленности, и прикрутить к ней AI, чем ждать, что чисто научная команда с нуля сделает отраслевой продукт».

Артур Самигуллин
Руководитель платформы Yandex AI Studio



«Свободные ниши находятся близко к производству – там, где необходимо работать с оборудованием, датчиками и сырыми данными, для которых нет готовых коннекторов. Продажи и закупки уже привлекают множество разработчиков, а интеграция производственных данных по-прежнему остаётся во многом незакрытым рынком, обладая при этом наибольшим потенциалом с точки зрения демонстрируемых бизнес-эффектов».

Алексей Бобок
СРО AI4BI



Медтех

~70 млрд ₽

Инвестиционный тезис:

Дефицит медицинского персонала, высокая нагрузка на врачей и большой объём неструктурированных клинических данных формируют спрос на AI-решения. Наиболее перспективны ассистенты для врача и пациента, решения в сфере анализа медицинских изображений. Устойчивое преимущество получают решения, интегрированные в МИС.

Финансы

~40 млрд ₽

Инвестиционный тезис:

Инвестиционная возможность связана с автоматизацией бухгалтерских процессов и аналитики: от обработки первичных документов и сверок до подготовки управленческой отчётности и интерпретации показателей. Такие продукты позволяют сократить объём ручной работы и упростить доступ к финансовым данным для специалистов без профильной подготовки. При этом высокие требования к точности ограничивают сценарии применения генеративного ИИ и требуют контроля результатов.

«В финансовой аналитике скорость внедрения генеративного ИИ будет определяться прежде всего требованиями к точности. Ошибки в расчётах и выводах быстро подрывают доверие к инструменту. Поэтому я ожидаю постепенного внедрения: сначала – помощь в поиске информации, подготовке комментариев и выполнении отдельных операций под контролем специалиста. Перспективную роль языковой модели я вижу в переводе запроса пользователя в команды, которые затем исполняются расчётными алгоритмами».

Роман Дидыч
Продуктовый директор
и Основатель Optimacros



Ключевые инвестиционно привлекательные сегменты российского рынка

Горизонтальный AI

Агенты и автоматизация
бизнес-процессов

~78 млрд ₽

Инвестиционный тезис:

Крупнейший и наиболее быстрорастущий горизонтальный сегмент. Рынок переходит от концепции «AI как помощник сотрудника» к автоматизации бизнес-процессов с помощью агентов, способных учитывать корпоративный контекст, принимать решения и выполнять многошаговые задачи, которые невозможно свести к линейному алгоритму. В результате ценность решения определяется уже не экономией отдельных часов, а сокращением объема ручного труда и автоматизацией целых функций.

Агенты могут стать новой организационной моделью, а не только инструментом экономии

«Автономные организации из AI-агентов — это уже не просто способ сократить расходы, а принципиально новая бизнес-модель. Агент-продакт, аналитик, разработчики, тестировщики и критики могут замкнуть полный цикл создания продукта и непрерывно работать 24/7. Такая система обеспечивает практически недостижимую для обычной команды скорость развития, более низкие операционные расходы и другой уровень качества и маржинальности бизнеса».

Армен Амирханян
Директор по развитию искусственного интеллекта Московской биржи



Продажи и клиентская
поддержка

~40 млрд ₽

Инвестиционный тезис:

Продажи и клиентская поддержка — одни из наиболее понятных направлений для расчёта экономического эффекта от AI: результат быстро отражается в росте конверсии и выручки, снижении стоимости контакта и нагрузки на операторов. Наибольший потенциал имеют решения, глубоко встроенные в CRM и процессы продаж. Регуляторный периметр мягкий — нет чувствительных данных уровня медицины/финансов в ядре продукта. Из рисков — низкий барьер входа и растущая взаимозаменяемость базовых чат-ботов.

«Продажи — одно из направлений, где эффект от ИИ можно связать непосредственно с деньгами. При этом задача не сводится к замене менеджера: доверие, коммуникация и ответственность остаются за человеком. ИИ может помогать находить информацию, готовить ответы и сопровождать отдельные этапы продажи. Для серьёзного корпоративного продукта потребуется постоянно обновляемая база знаний о продуктах и специфике компании — и инструменты, позволяющие сотрудникам поддерживать её актуальность».

Роман Дидыч
Продуктовый директор и Основатель Optimacros



Управление
цепочками поставок

~20 млрд ₽

Инвестиционный тезис:

AI в управлении цепочками поставок позволяет оптимизировать закупки, сокращать избыточные запасы и снижать потери от дефицита товаров и сырья. Влияние на издержки и оборотный капитал позволяет оценивать эффект внедрения в финансовых показателях. Для независимых российских компаний перспективны специализированные решения для прогнозирования спроса, планирования поставок и закупочной аналитики. Ключевой барьер входа — доступ к данным заказчиков. Закупочные цены, объёмы и условия контрактов относятся к чувствительной коммерческой информации: эти сведения, как правило, отсутствуют в открытом доступе, а крупные холдинги неохотно передают их сторонним разработчикам.

«Прикладной AI в закупках начинается не с модных агентов, а с довольно скучной работы с данными: справочниками учётных систем, нормализацией номенклатуры и работой с историей закупок. <...> Можно сколько угодно демонстрировать красивую AI-технологию, но если система не прошла путь эволюции и обучения на реальных данных (зачастую грязных, задублированных или не связанных), реальной ценности эта технология не даст. Главный барьер для AI-компаний и стартапов в закупках — доступ к большим массивам реальных данных холдингов. Крупный индустриал (промышленные предприятия) воспринимает закупочные и ценовые данные как сверхчувствительную информацию, и доступ новым игрокам рынка к ней, по сути, закрыт. Без этого обучить свои модели и доказать результат реальными кейсами — невозможно. <...>. Очевидно, что рынок компаний, представляющих AI-решения в закупках, быстро разделится: буквально несколько игроков с доступом к данным и сильными кейсами, и все остальные, которым останется встраивать свои технологии в существующие закупочные экосистемы».

Юрий Логвинов
Директор по продукту Uptrade (ООО «Аптрейд»)



Ключевые инвестиционно привлекательные сегменты российского рынка

Инфраструктура

Развертывание, оркестрация и оптимизация инференса

~27 млрд ₽

Инвестиционный тезис:

Рост числа моделей и дорогие вычисления создают спрос на платформы, которые позволяют разворачивать AI в корпоративном контуре, маршрутизировать запросы между моделями, балансировать качество и цену ответа, управлять нагрузкой и снижать затраты на инференс. Для российского рынка это особенно актуально из-за распространённости on-premise-развёртываний и ограниченного доступа к GPU. Однако open source быстро снижает различия между базовыми инструментами.

Оптимизация inference открывает инфраструктурную нишу с измеримым эффектом

«Сейчас перспективнее фокусироваться на оптимизации инференса — кэшировании, сжатии контекста и сокращении избыточного расхода токенов. Такие решения уже позволяют снижать потребление токенов на 60–80% и одновременно повышать качество работы агентов: чем меньше лишней информации поступает в модель, тем лучше она удерживает внимание на значимых данных. При этом индустрия развивается настолько быстро, что оптимизацией пока почти никто системно не занимается, а значит, здесь сохраняется большое пространство для появления новых компаний».

Армен Амирханян
Директор по развитию искусственного интеллекта Московской биржи



Инфраструктура данных

~22 млрд ₽

Инвестиционный тезис:

Дефицит доступных качественных данных становится одним из ключевых ограничений развития AI. Качество AI-решений напрямую зависит от качества и доступности данных, тогда как в российских компаниях данные часто разрознены между legacy-системами и не готовы к использованию моделями.

Другие сегменты

Physical AI

~25 млрд ₽

Инвестиционный тезис:

Physical AI — следующий этап развития AI, при котором модели переходят от обработки цифровой информации к управлению объектами в физическом мире. Наибольший потенциал сосредоточен в робототехнике, дронах, беспилотном транспорте и автономной сельхозтехнике. В робототехнике необходимость адаптации AI под конкретное оборудование, среду и бизнес-процесс создаёт высокий технологический барьер и формирует ёмкий рынок специализированных решений. Для России также перспективна автономная сельхозтехника благодаря масштабу аграрного сектора. Однако как технология Physical AI пока находится на ранней стадии развития.

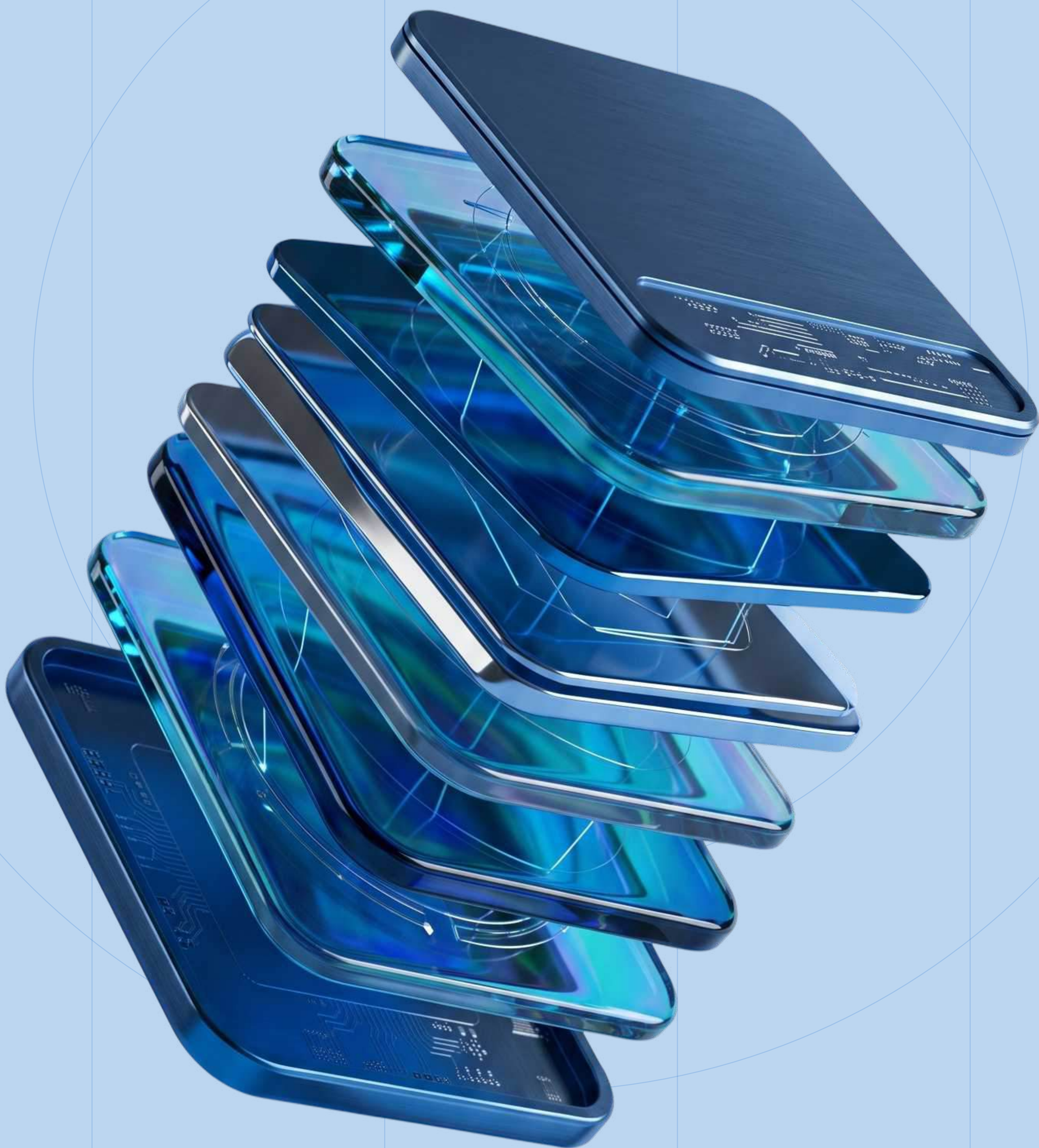
Услуги

~35 млрд ₽

Инвестиционный тезис:

AI становится обязательной частью корпоративной повестки, но большинству компаний не хватает внутренней экспертизы для его внедрения. Поскольку AI-решения редко работают «из коробки», растёт спрос на подрядчиков, которые адаптируют модели под данные, процессы и IT-контур конкретного заказчика. Отдельный рынок формируют консалтинг по стратегии AI-трансформации и обучение сотрудников работе с AI-системами.

Специфика и барьеры развития российского рынка



Специфика и барьеры развития российского рынка

Российская специфика внедрения: GenAI быстрее всего внедряется в линейных бизнес-процессах бэк-офиса и в разработке

GenAI активно внедряется

01

Разработка кода

Сопilot-решения и ассистенты для разработчиков

02

HR и рекрутмент

Скоринг резюме, коммуникации, автоматизация рутинных этапов воронки

03

Корпоративный поиск

AI-ассистент для поиска по внутренним документам, регламентам, базам знаний

04

Клиентский сервис

Чат-боты для коммуникации с клиентами

Большой объём накопленных данных, лёгкость для внедрения, быстрый и измеримый эффект

«Основные успешные кейсы AI сегодня лежат в автоматизации линейных бизнес-функций: HR, юридических процессов и закупок – от отбора резюме и анализа нормативной базы до подготовки закупочной документации».

Андрей Фиохин
Генеральный директор
ООО «Айтисфера»
(входит в ГК «АЗОТ»)

«Больших кейсов, где AI действительно трансформировал компанию, я пока не видел. В основном это точечные решения или развитие сценариев цифровизации, которые существовали и раньше».

Антон Шмаков
Технический директор,
ПАО Группа Астра

GenAI внедряется медленно

01

Промышленность

02

Строительство

03

Медицина

04

Нефть и газ

05

Металлургия и горная добыча

06

Сельское хозяйство

Высокая цена ошибки, высокий CapEx, регуляторные барьеры, слабая база данных

«По моим наблюдениям, интерес российских компаний к ИИ пока заметно опережает готовность покупать решения. Демонстрация прототипа вызывает любопытство, но затем заказчик спрашивает: где реальные внедрения, как решены вопросы безопасности и какова окупаемость? Для перехода к масштабному спросу нужны продукты, встроенные в корпоративные процессы, и подтверждённые примеры экономического эффекта. Одной убедительной демонстрации возможностей модели недостаточно».

Роман Дидыч
Продуктовый директор
и Основатель Optimacros

Факторы, определяющие скорость внедрения

01

Низкая цена ошибки ускоряет внедрение

Чем ниже риск критичных последствий от ошибки AI, тем проще запустить пилот и масштабировать решение

02

Простая интеграция в существующий бизнес-процесс

AI быстрее внедряется там, где не требуется глубокая перестройка IT-архитектуры, регламентов и ролей сотрудников.

03

Наличие оцифрованных данных

«Российские компании одновременно проходят через импортозамещение, незавершенную цифровизацию и внедрение AI. Особенно в тяжёлых отраслях до сих пор сохраняются бумага, ручной труд и слабо формализованные процессы. AI не может компенсировать отсутствие базовой цифровизации: если процесс не формализован, а данные не структурированы, встраивать модель фактически некуда – поэтому многие эксперименты пока не дают значимого эффекта».

Антон Шмаков
Технический директор,
ПАО Группа Астра

04

Быстрый и измеримый эффект для бизнеса

Эффект виден в коротком горизонте: снижение затрат, ускорение процессов, повышение качества.

«В промышленности управленческие решения часто принимают руководители-прагматики с инженерным прошлым, которым нужен подтверждённый эффект: покажите, где это уже работает у сопоставимой компании и какие результаты были получены. Одних обещаний недостаточно – рынку не хватает конкретных кейсов с доказанным экономическим эффектом».

Алексей Бобок
СРО AI4BI

Барьеры развития российского AI-рынка: основные выводы

01. Оборудование: сложность закупки и высокая стоимость GPU

В 2022 году NVIDIA приостановила прямые продажи в Россию и позднее закрыла российские операции. В августе того же года США ввели лицензионные ограничения на экспорт в Россию передовых ускорителей NVIDIA A100 и H100 и систем на их основе, в результате чего российский рынок оказался зависим от параллельного импорта, ведущего к удорожанию оборудования.

«Если раскладывать рынок, то всё начинается с инфраструктуры, с которой у нас большие проблемы. Нужны GPU, GPU дорогие, пути и способы их доставки тернистые».

Лариса Малькова
Управляющий директор,
Руководитель практики
«Прикладной Искусственный
Интеллект», Ахенix

«Все AI-стартапы хотят получить токены или GPU — многие готовы брать вычислительные мощности вместо денег».

Игорь Маслов
Главный технический
директор Т-Банка

«В развитии генеративного ИИ ключевым фактором являются не столько математические архитектуры, сколько доступ к вычислительным мощностям. Сегодня в России нет той GPU-инфраструктуры, которая создаётся в США и Китае».

В таких условиях практически невозможно самостоятельно преодолеть разрыв в базовых моделях, поэтому наиболее реалистичная стратегия для российских компаний — искать узкоспециализированные прикладные сценарии, в которых можно создавать измеримый бизнес-эффект».

Армен Амирханян
Директор по развитию искусственного
интеллекта Московской биржи

Корпоративные участники рынка также отмечают сложность обоснования инвестиций в вычислительные мощности. IT-подразделения видят потребность в закупке GPU, однако необходимый объём мощностей сложно заранее связать с будущей загрузкой инфраструктуры и измеримым финансовым эффектом от AI-проектов. В результате для менеджмента такие инвестиции выглядят как значительный CapEx с неочевидной окупаемостью.

02. Требования к on-prem развертыванию и безопасности данных

Требования к локальному хранению и защите данных заставляют многие компании разворачивать AI-системы внутри корпоративного контура. Для быстрого масштабирования таких решений по-прежнему важна облачная инфраструктура.

По данным «Яков и Партнёры», высокий уровень внедрения AI характерен прежде всего для компаний, активно использующих public cloud, — их доля составляет 30%. В то же время при низком уровне использования облачных сервисов этот показатель не превышает 2%.

«Наблюдается выраженная положительная корреляция: чем выше зрелость облачных решений в компании, тем успешнее она внедряет технологии искусственного интеллекта. Это связано с тем, что развитая облачная инфраструктура обеспечивает масштабируемость, гибкость и вычислительные ресурсы, необходимые для эффективного использования ИИ».

Илья Голод
Руководитель проектов «Яков и Партнёры»
Пресс-релиз «Яков и Партнёры», 24 ноября 2025 года

Российский AI-рынок сталкивается со структурным конфликтом между безопасностью данных и скоростью технологического развития.

«На российском рынке технологически будущее принадлежит SaaS и гибридным решениям, но регуляторные ограничения продолжают возвращать компании к on-premise. Это замедляет пилотирование AI-решений и вынуждает промышленные предприятия разворачивать дорогостоящую инфраструктуру в закрытом контуре. Необходимо найти баланс между защитой критически важных данных и технологическим развитием промышленности — обеспечить хранение чувствительной информации внутри России, сохранив возможность использовать защищённую облачную инфраструктуру».

Александр Вибе
Директор Центра ИИ-решений для B2B Сбербанка

03. Незрелость рынка

Компании пока не готовы перестраивать процессы под AI, а у пользователей не сформировалась привычка регулярно использовать GenAI-продукты. Для большинства LLM остаются инструментом для отдельных задач, а не частью повседневной практики.

«Три ключевых препятствия: 1) Сложность покупки и высокая стоимость оборудования; 2) Низкая зрелость экосистемы в России (нет аналогов ChatGPT и пр.), из-за чего нет потребительской привычки; 3) Необходимость перестройки процессов под AI, а не навязывание технологий сверху. Последний пункт особенно важен для корпоративных внедрений».

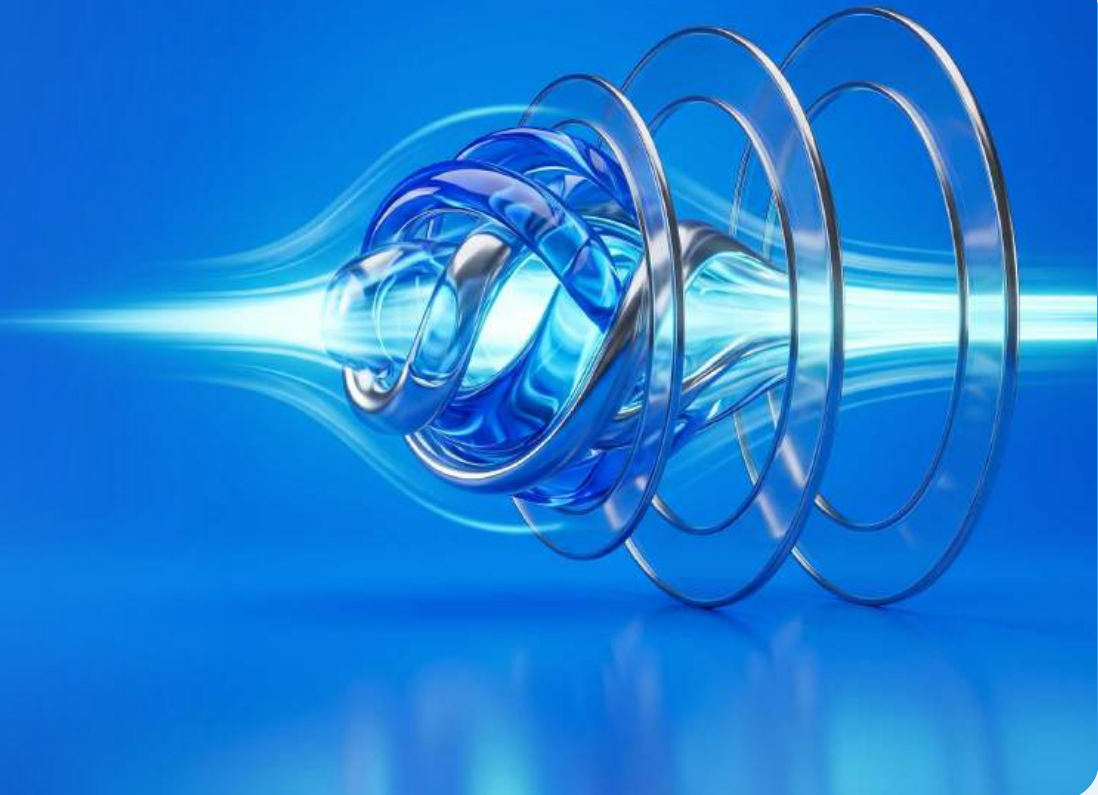
Андрей Рыбинцев
Управляющий директор по ИИ Авито

04. Технологическая изоляция от глобальной AI-экосистемы

Россия отсутствует в официальных списках поддерживаемых стран для API OpenAI и Anthropic, а получение доступа может привести к блокировке или приостановке аккаунта. Ограничен доступ к крупнейшим глобальным облачным платформам – в марте 2022 года AWS прекратил регистрацию новых клиентов из России и Белоруссии. Microsoft одновременно приостановила все новые продажи своих продуктов и сервисов в России, что затрагивает в том числе Azure и связанные с ним инструменты.

Основные выводы

Российский рынок AI формируется на пересечении двух разнонаправленных трендов: растущего корпоративного спроса на внедрение AI и ограниченного доступа независимых компаний к капиталу, данным и вычислительной инфраструктуре.



01

AI в России развивается по корпоративной модели

Развитие фундаментальных моделей и вычислительной инфраструктуры сосредоточено преимущественно внутри крупных технологических экосистем. Публично заявленные вложения BigTech в AI в 52 раза превышают венчурное финансирование AI-стартапов. Дефицит венчурного капитала сдерживает появление независимых AI-стартапов, и вместе с ним — спрос на инфраструктуру.

32%

Доля вложений в AI-стартапы на российском венчурном рынке в 2025 г.

02

На маржинальность AI-бизнеса влияет стоимость внедрения и затраты на инференс

Расходы на инференс растут вместе с использованием, а адаптация под каждого заказчика ограничивает масштабируемость бизнеса. Для инвестора принципиальны экономика промышленной эксплуатации, возможность повторных внедрений и доля продуктовой выручки: высокая зависимость от кастомной разработки влияет на маржинальность и оценку компании.

0,0019%

ВВП России пришлось на венчурные инвестиции в AI в 2025 г.

03

Российские AI-стартапы ограничены в доступе к капиталу, вычислительным мощностям и данным

Потенциальная модель взаимодействия для стартапов — партнёрство, где BigTech даёт инфраструктуру, корпоративный заказчик — данные и контур для внедрения, а стартап создаёт специализированный продукт, опираясь на отраслевую экспертизу, скорость и гибкость.

84 млрд ₽

Дополнительного венчурного капитала ежегодно необходимо для развития независимого AI-рынка в России

04

Главный конкурент AI-стартапа в России — IT-отдел внутри компании-заказчика

GenAI кратно удешевил и ускорил скорость разработки прототипа, и Build vs Buy решается в пользу Build. По мере роста доступности базовых моделей устойчивое преимущество для AI-стартапа создают отраслевая экспертиза, данные и интеграция в процессы заказчика. Независимая компания должна подтверждать экономический эффект и обладать компетенциями или активами, которые делают покупку её решения привлекательнее собственной разработки.

52x

На каждый доллар венчурных инвестиций в российские AI-стартапы приходится 52 доллара инвестиций BigTech-корпораций в ИИ

05

Рациональная стратегия для частного инвестора — не конкурировать с BigTech

Фундаментальные модели становятся взаимозаменяемыми, а их разработка, «железо» и облака требуют больших инвестиций в CapEx и длинного горизонта инвестирования. Возможности для частного инвестора в России — в прикладных продуктах.

06

Инвестиционно привлекательными выглядят сегменты с высоким/средним барьером входа, низким CapEx и измеримым бизнес-эффектом:

- Вертикальные решения в сфере медицины, промышленности и финансов;
- Горизонтальные решения в сфере автоматизации продаж и агенты;
- Точечные решения в сфере данных и связующего ПО;
- Услуги и Physical AI.

561 млрд ₽

Совокупный потенциал российского AI-рынка, доступного независимым частным игрокам к 2030 году

Благодарности

Команда KAMA FLOW благодарит экспертов, которые поделились своим профессиональным мнением и опытом при подготовке исследования. Их оценки и комментарии позволили дополнить анализ рынка практическим взглядом на происходящие в отрасли изменения и перспективы ее развития.



Игорь Маслов
Главный технический директор Т-Банка



Александр Вибе
Директор Центра ИИ-решений для B2B СберБанка



Андрей Рыбинцев
Управляющий директор по ИИ Авито



Артур Самигуллин
Руководитель платформы Yandex AI Studio



Алексей Бобок
СРО AI4BI



Антон Шмаков
Технический директор ПАО «Группа Астра»



Армен Амирханян
Директор по развитию искусственного интеллекта Московской биржи



Алексей Постернак
Директор по инвестициям и развитию MWS



Лариса Малькова
Управляющий директор, руководитель практики «Прикладной искусственный интеллект» Ahenix



Евгений Кокуйкин
CEO HiveTrace



Виктор Поздняков
Руководитель направления стратегических инвестиций Т-Банка



Анна Буракова
Руководитель M&A IBS



Андрей Фиохи́н
Генеральный директор ООО «Айтисфера» (ГК «Азот»)



Роман Дидыч
Продуктовый директор и Основатель Optimacros



Юрий Логвинов
Директор по продукту Uprade



Связаться с авторами и запросить комментарии:

info@kamaflow.com pr@kamaflow.com



Адрес

Москва
Малый Ивановский переулок,
7-9 ст1, 1 этаж, офис 2/1

Новосибирск
ул. Николаева 11/5,
офис 802

