

# Генеративные нейросети

Восстание машин  
или новая экономика

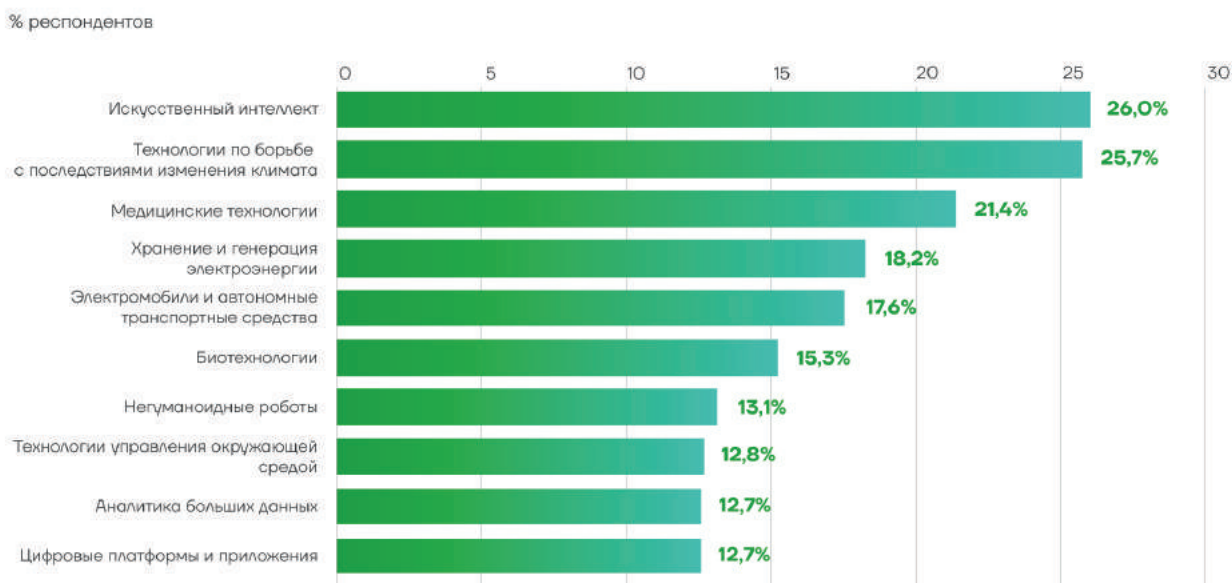
# Генеративные нейросети: восстание машин или новая экономика

- ▶ Генеративные системы искусственного интеллекта (ИИ), которые способны создавать изображения и тексты, спровоцировали новую волну интереса к этой технологии. Бесплатный чат-бот ChatGPT набрал 100 миллионов пользователей меньше чем за три месяца.
- ▶ Для того, чтобы довести технологию до состояния, когда она сможет серьезно улучшить производительность труда, потребуются многомиллиардные инвестиции и дополнительное развитие вычислительных мощностей.
- ▶ Проблемой генеративных систем ИИ стала возможность использования технологии в преступных целях — для создания спама, дипфейков и фишинга.
- ▶ Формирование законодательства и доступность вычислительных мощностей вкупе с данными для обучения определит экономическое будущее генеративных систем.

## Радужные прогнозы

По результатам опроса, проведенного среди 12 тысяч участников Всемирного экономического форума в Давосе в 2023 году, технологии, связанные с ИИ, имеют стратегическую важность, наряду с инновациями в борьбе с последствиями изменения климата и медицинскими ноу-хау.

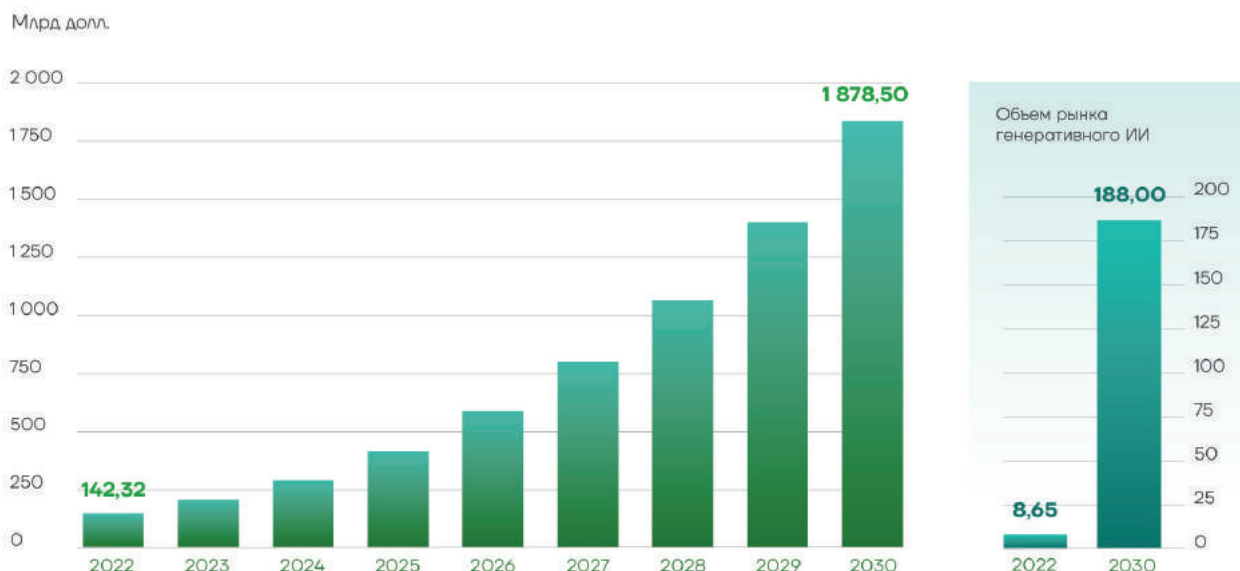
### Технологии стратегической важности



Источник данных: WEF Markets of Tomorrow 2023

По прогнозам аналитиков Next Move Strategy Consulting, до конца текущего десятилетия объем рынка продуктов, связанных с ИИ, вырастет почти в 10 раз и приблизится к 2 трлн долл. Самое большое распространение они получат в таких отраслях как управление цепочками поставок, маркетинг, дизайн продуктов, аналитика данных.

## Глобальный рынок ИИ до 2030 года



Источник данных: Next Move Strategy Consulting

Самым быстрорастущим сегментом рынка ИИ станут генеративные модели. По подсчетам Brainy Insights, за ближайшие 10 лет этот сегмент увеличится в 20 раз — с 8,65 млрд долл. в 2022 году до 188 млрд долл. к 2032 году. В отличие от прочих систем машинного обучения, генеративные модели обладают широкими возможностями создания новых текстов и изображений, основываясь на закономерностях, которые были установлены из массивов уже существующих данных.

Эти модели благодаря быстрому росту доступной вычислительной мощности и появлению больших наборов данных за последние пять лет сделали серьезный скачок в решении задач, которые раньше считались неразрешимыми для машин.

## Нейросети-трансформеры

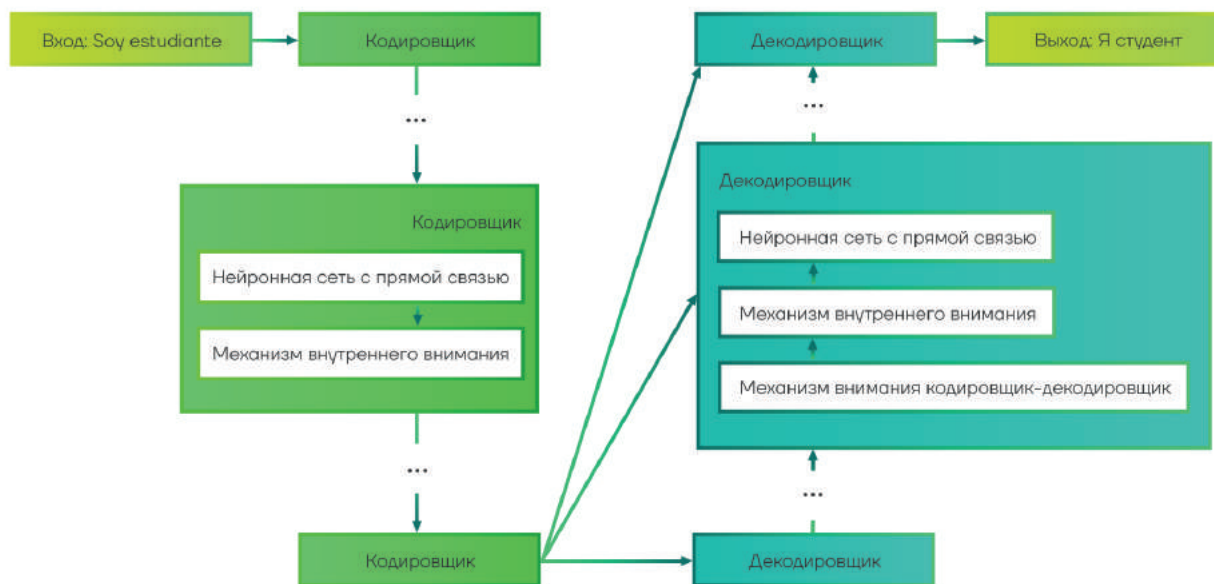
Еще одним фактором, который обусловил качественный рост возможностей ИИ, стало появление в 2017 году новой архитектуры моделей ИИ — трансформеров.

Основная идея transformer-модели заключается в использовании механизма внимания (attention) для обработки последовательностей данных. Вместо того, чтобы один за другим обрабатывать каждый элемент предложенного текста,

модель может одновременно «смотреть» на все элементы и определить, какие из них наиболее важны для решения задачи.

Трансформер можно сравнить с учителем, который может одновременно обращать внимание на всех учеников в классе и определять, кому нужно уделить больше времени и усилий, чтобы помогать отстающим лучше усвоить материал.

### Принцип работы модели Transformer

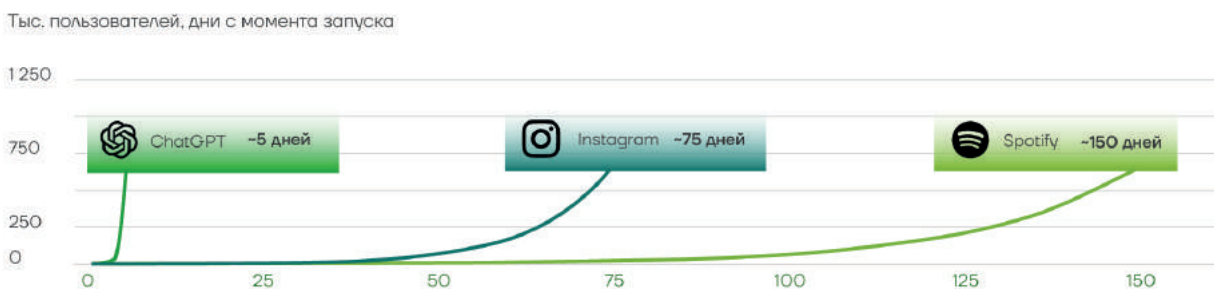


Источник данных: The Illustrated Transformer

Трансформер обычно состоит из двух компонентов: кодировщика и декодировщика. Кодировщик принимает на вход последовательность данных и преобразует ее во внутреннее представление, которое может быть использовано декодировщиком для генерации выходной последовательности. Например, перевода на другой язык или написания текста на заданную тематику.

Эта модель лежит в основе многих современных алгоритмов обработки естественного языка. Она была использована компанией OpenAI для создания языковой модели GPT-3. Чат-бот ChatGPT, созданный с ее помощью, наделал много шума в новостях.

### Скорость привлечения 1 млн пользователей

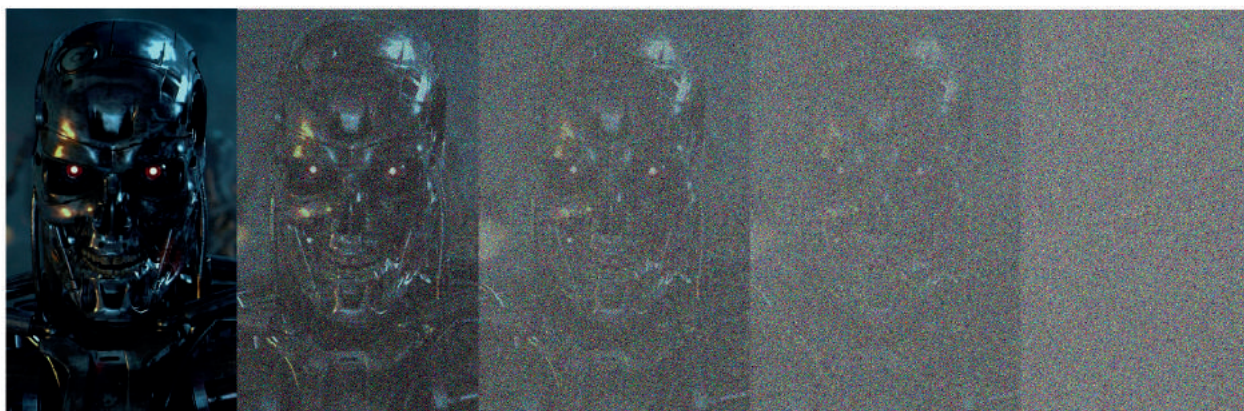


Источник данных: Google

# Диффузионная модель

## Диффузионная модель ИИ

Прямая диффузия (от данных к шуму)



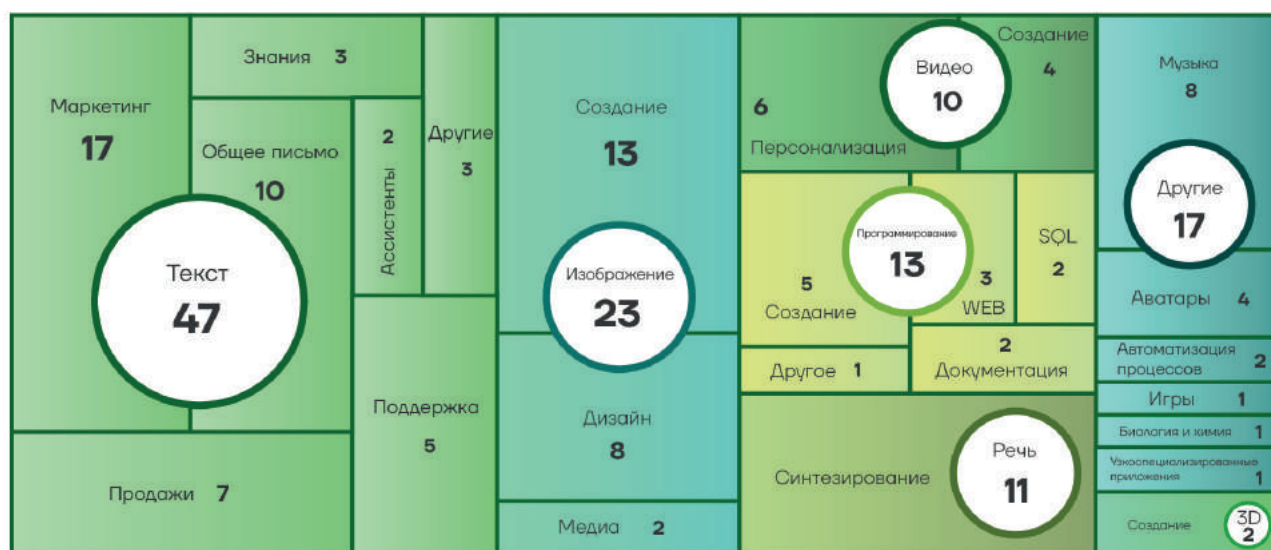
Обратная диффузия (от шума к данным)

Источник данных: Score-Based Generative Modeling through Stochastic Differential Equations

Для работы со звуком и изображениями используется класс диффузионных (diffusion) моделей. Самые популярные методы их создания описаны в работе 2020 года. Диффузоры устроены так, что в обучающие изображения постепенно вносится шум, который шаг за шагом стирает детали. Затем модель подбирает параметры так, чтобы из шума восстановить образец.

## Сегменты применения ИИ

Количество компаний



На практике для решения задачи генерации изображения по текстовому запросу используется несколько моделей. Например, CLIP от OpenAI. Ее обучали на большой выборке пар «изображение-описание» переводить текст в картинку с низким разрешением. Затем диффузионная модель увеличивает ее и дополняет деталями.

На начало 2023 года различные системы генеративного ИИ стали выходить из узкой академической среды и оформляться в конкретные приложения.

## Работа с текстом

Самый развитый сектор — работа с текстом. Уже существующие языковые модели удовлетворительно справляются с такими задачами как написание писем в заданном стиле и маркетинговых материалов, реферирование текста, составление ответов на типичные обращения в службы клиентской поддержки.

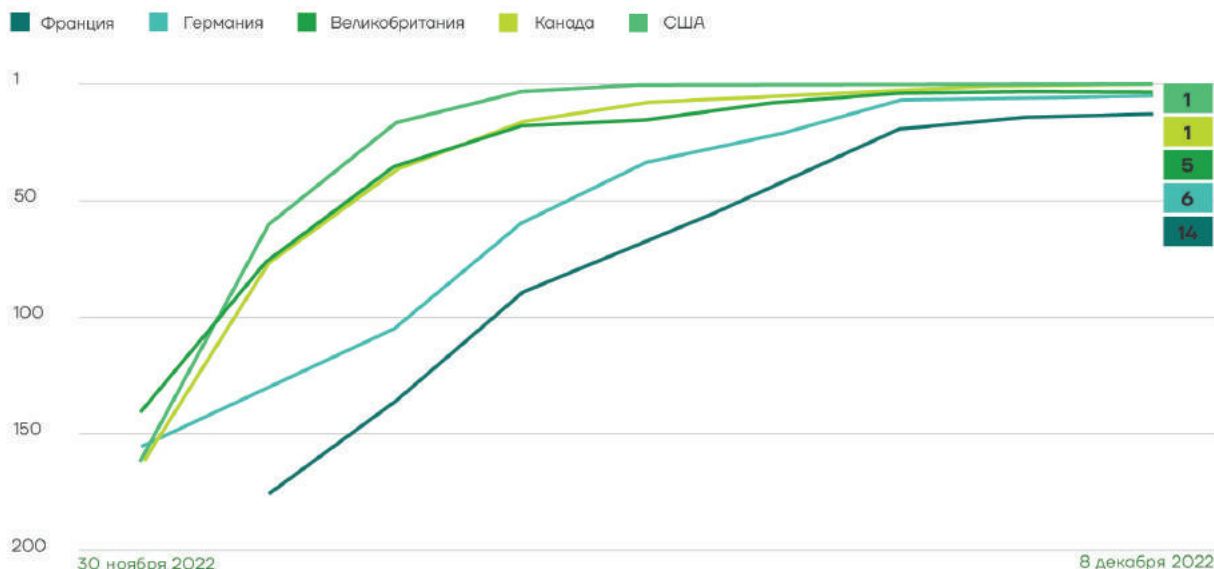
Генерация кода по заданным параметрам, документация программ, превращение запросов на естественном языке в SQL-код для работы с базами данных являются хорошо структурированными задачами, которые лучше других подходят для решения с помощью систем машинного обучения. Бесплатный чат-бот ChatGPT набрал 100 миллионов пользователей меньше чем за три месяца.

Большие языковые генеративные модели создают и в России. В июне 2022 года Яндекс открыл для разработчиков модель YaLM100B. Для ее обучения было использовано 1,7 ТБ текстов, а сам процесс подбора оптимальных значений 100 млрд параметров занял 65 дней, при этом использовалось 800 графических ускорителей A100.

## ИИ-художник

Формирование новых изображений находит своё применение в рекламе, дизайне новых товаров, интерьеров и различных макетов. Генерация аватаров и других изображений для пользователей уже стала востребованной услугой с высоким коммерческим потенциалом. Приложение Lensa AI, разработанное бывшими сотрудниками Яндекса, за 10 дней в начале декабря 2022 года вышло на первое место по выручке в США и Канаде.

## Позиция Lensa AI в рейтинге выручки приложений в App Store



Источник данных: SensorTower

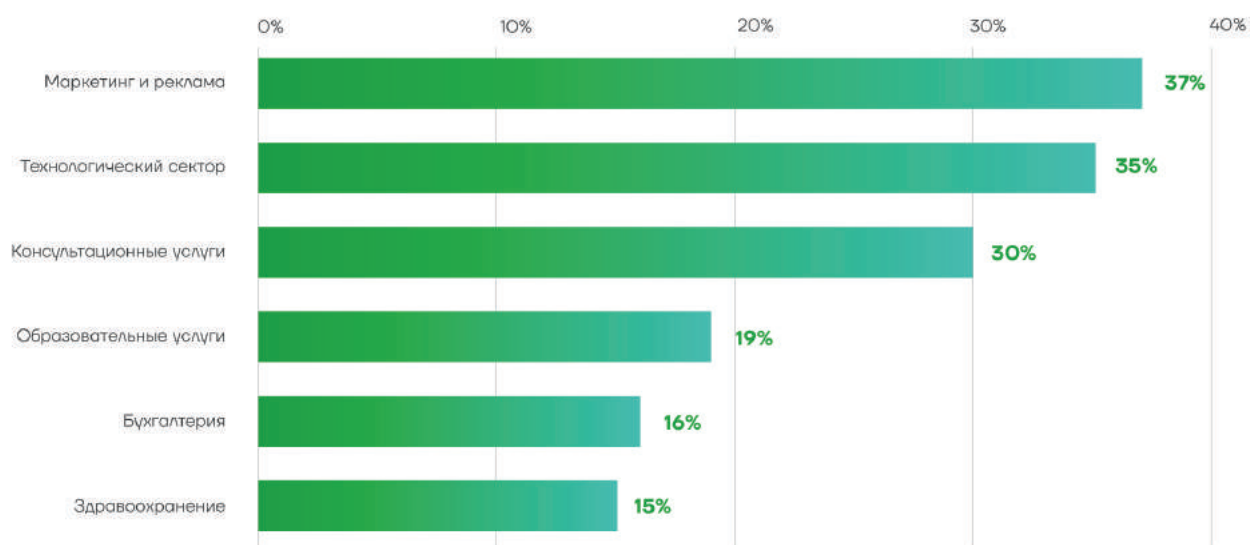
В ноябре 2023 года исследователи Sber AI представили диффузионную модель Kandinsky 2.0, обученную на объединенном датасете Sber AI и SberDevices из 1 млрд пар «текст — изображение».

По схожему сценарию, но с отставанием на пару лет, развивается ситуация с генерацией видео и созданием озвучки. На начало 2023 года качество недостаточно высокое, чтобы использовать ИИ без дополнительной обработки.

## Ресурсы для большого скачка

### Внедрение генеративных ИИ на рабочих местах в США

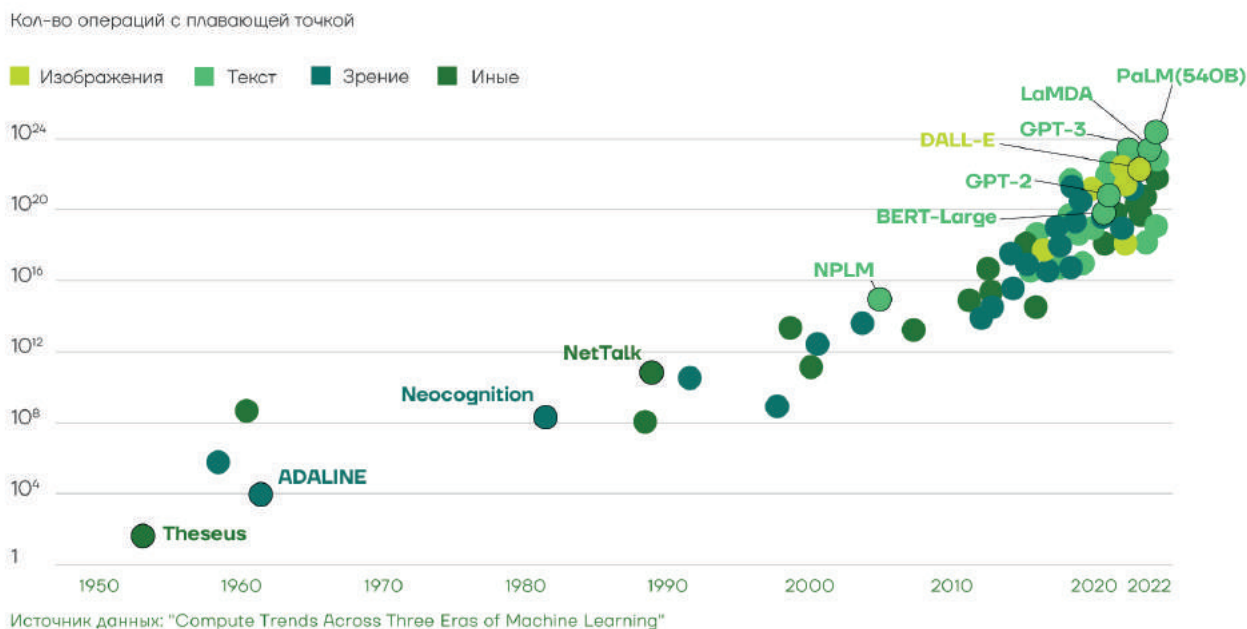
2022, по отраслям



Источник данных: Fishbowl

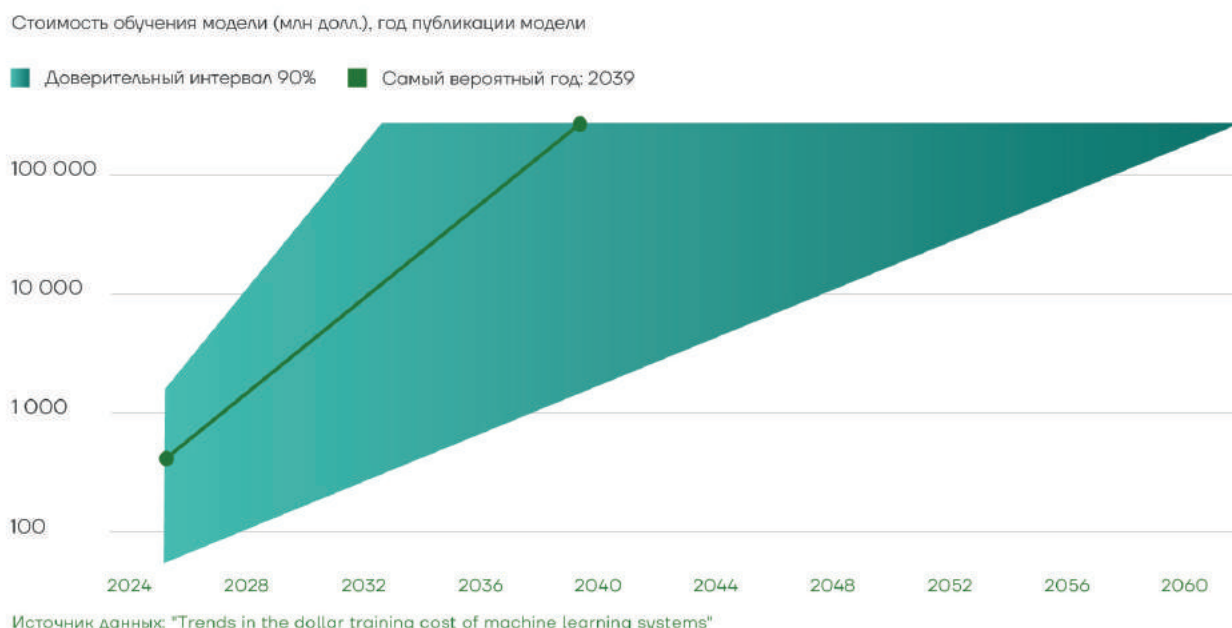
Использование генеративных моделей ИИ в бизнесе быстро становится распространенной практикой. По результатам опроса, который провела компания Fishbowl, в США каждая третья компания из сектора рекламы и маркетинга в 2023 году планирует использовать этот инструмент в своей работе. Схожие результаты в технологическом секторе и сфере консультационных услуг.

### Вычислительная стоимость обучения моделей



Однако все современные подходы к созданию моделей ИИ требуют колоссальных объемов вычислительной мощности. Если до 2010 года требования к объемам вычислений росли в примерном соответствии с законом Мура (мощности удваивались каждые полтора года), то с появлением больших моделей темп возрос. С 2010 по 2022 год объем вычислений, необходимых для обучения одной модели увеличился в 10 миллиардов раз.

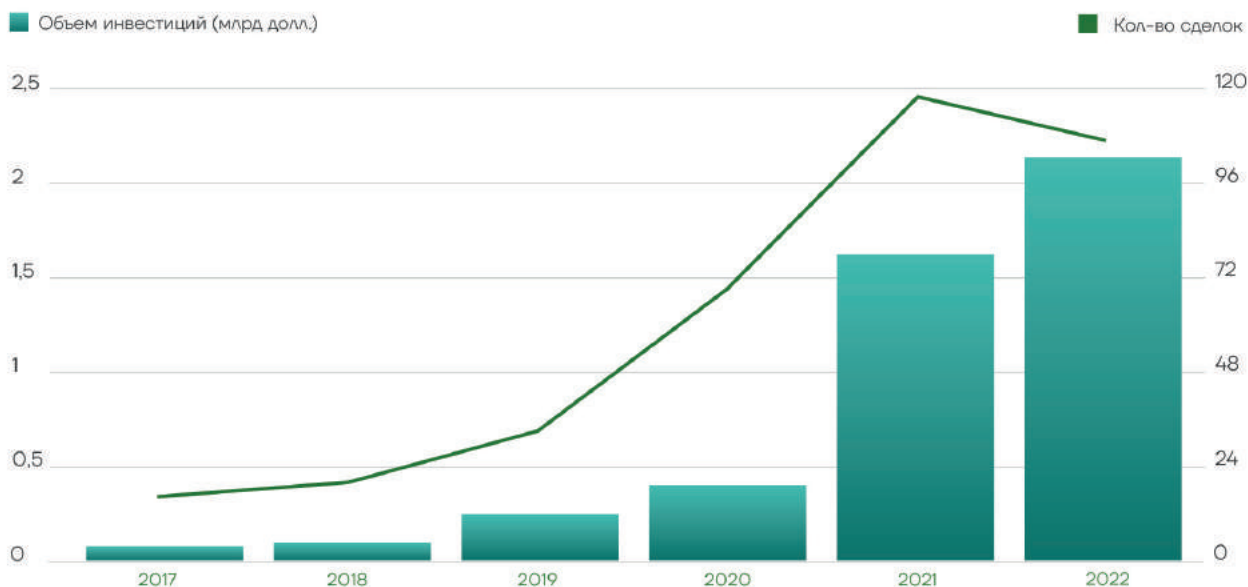
### Когда расходы на обучение 1 модели достигнут 1% ВВП США?





Если продлить этот тренд в будущее, то на горизонте двух десятилетий стоимость сбора датасета для обучения, подготовки данных и, непосредственно, самого обучения одной большой языковой модели превысит 200 млрд долл., то есть составит около 1% ВВП США за 2021 год. Бен Котье (Ben Cottier) из исследовательского института EpochAI закладывает в свой прогноз большую долю неопределенности, но не сомневается, что в ближайшем будущем создание и использование больших языковых моделей будет доступно только крупным корпорациям и правительствам государств с развитой экономикой.

## Объем венчурных инвестиций в генеративный ИИ



Источник данных: PitchBook

## Восток vs Запад

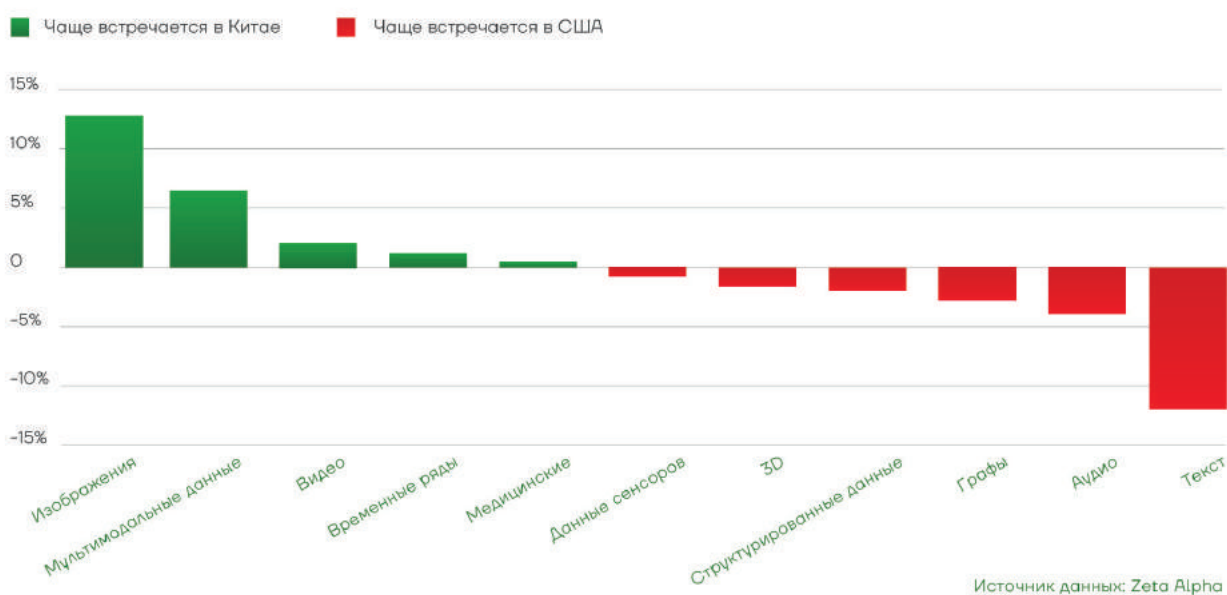
Огромная стоимость создания моделей вынуждает стартапы в области генеративного ИИ активно привлекать финансирование. За 2022 год объем венчурных инвестиций превысил 2 млрд долл., и это в три с лишним раза больше, чем в 2020-м. В 2023 приток средств стал еще более щедрым, при этом инвестиции идут в основном на американский рынок. К середине марта 2023 года о крупных сделках в этой сфере объявили Character.ai, получивший от фонда Andreessen Horowitz 200 млн долл. при оценке всей компании в 1 млрд долл.; OpenAI, заключивший многолетний договор о финансировании с Microsoft на 10 млрд долл. при оценке всей компании в 29 млрд долл.; Anthropic, получивший от Google 300 млн долл. при оценке около 3 млрд долл..

В начале марта Salesforce, которая разрабатывает системы автоматизированного взаимодействия с клиентами, объявила о запуске разработанной вместе с инженерами из OpenAI модели Einstein GPT. Она может составлять персонализированные рассылки по электронной почте, таргетированный контент

для менеджеров по продажам и отвечать на запросы в службу поддержки. Чтобы в будущем не полагаться только на одного поставщика ИИ-технологий, Salesforce создала специальный фонд в размере 250 млн долл., который будет инвестировать в перспективные стартапы в этой области.

По оценкам Grand View Research, объем инвестиций в ИИ на конец 2022 года в США составил 41% от общемирового. Однако это агентство прогнозирует самые быстрые темпы роста рынка ИИ, включая генеративные модели, в азиатско-тихоокеанском регионе.

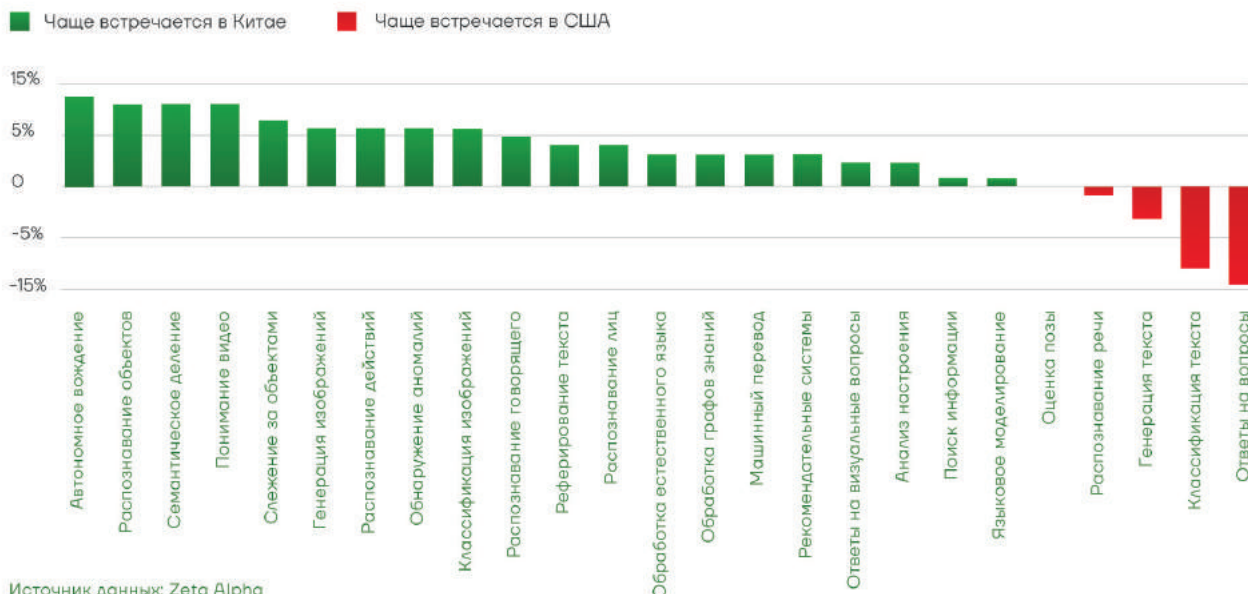
### Тип данных, фигурирующий в научных публикациях по теме ИИ в КНР и США



Сектор ИИ в Китае делает упор на другие типы задач. Среди опубликованных исследований в этой сфере более существенный акцент, по сравнению с США, делается на применении ИИ в области автономного движения, распознавания и слежения за объектами и ориентирования в пространстве. Соответственно, в исследованиях чаще фигурируют методы обработки изображений и видео, и гораздо реже — задачи, связанные с текстами и аудио.

Темпы роста количества публикаций на тему ИИ в международных базах знаний от исследовательских групп, расположенных в Китае, в 2022 были вдвое выше, чем от американских и европейских коллег. Но по общим объемам китайцы пока отстают. По подсчетам инвестиционного фонда Zeta Alpha в 2022 году они опубликовали 7,3 тысячи работ против 11,2 тыс у США. Однако отставание тут может быть просто ошибкой перевода. Если к опубликованным в международных журналах работам прибавить те, которые выходили только на китайском языке и находятся исключительно в национальной базе знаний КНР, получается, что за последние 12 лет объем публикаций по теме ИИ от китайских ученых превышает американский в 4,5 раза.

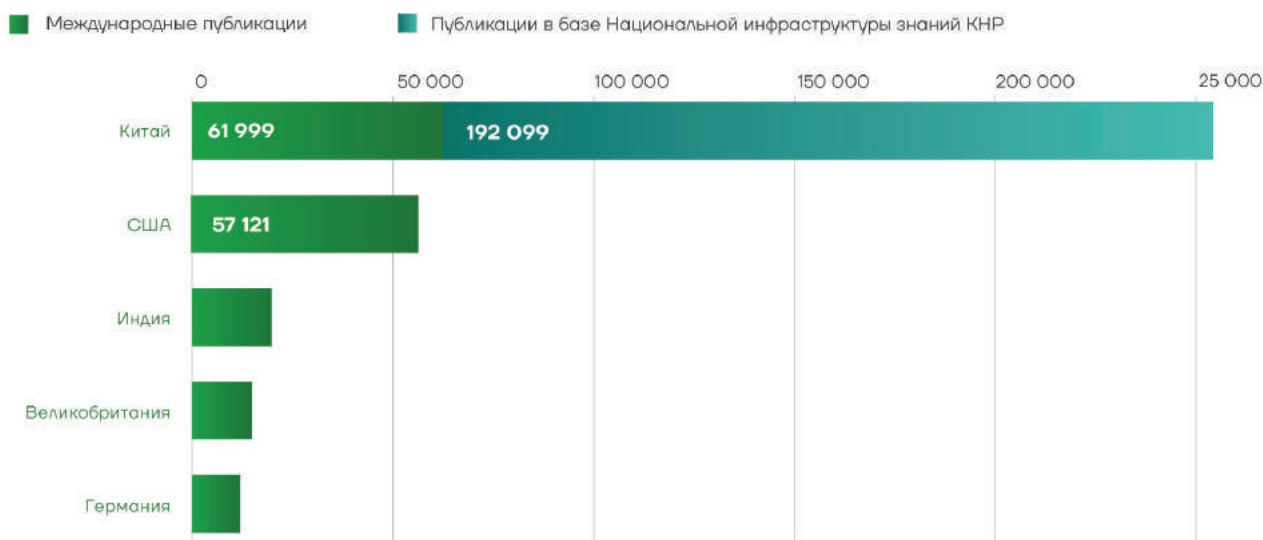
## Задачи машинного обучения в научных публикациях по теме ИИ в КНР и США



Дальнейшее развитие ИИ в Китае будет зависеть от доступа к мощному вычислительному оборудованию. В первую очередь, к ускорителям с характеристиками, сравнимыми с H100 от Nvidia, импорт которых в КНР был ограничен в 2022 году.

## Количество публикаций по теме ИИ

2010-2022

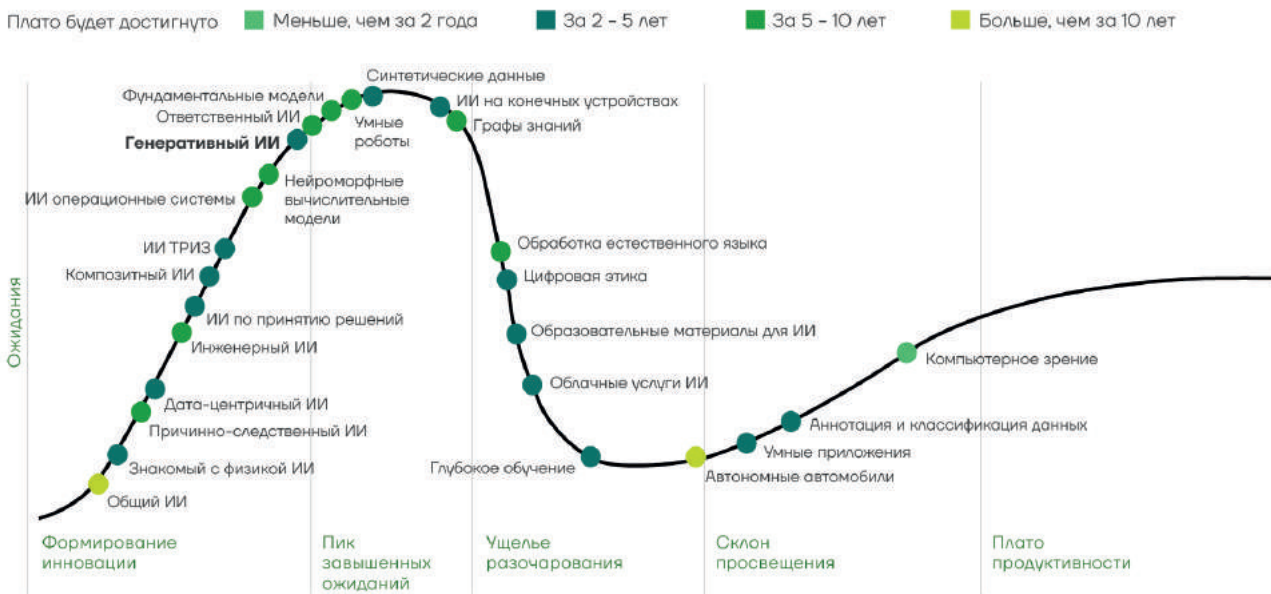


## Как далеко до плато продуктивности?

Появление генеративных моделей ИИ, способных создавать качественные изображения и связный текст, подстегнуло интерес широкой публики и инвесторов к этому сектору. Вопрос о том, насколько большие языковые модели смогут улучшить производительность офисного и творческого труда, остается открытым.

Сгенерированные изображения нередко выглядят абсурдно, а выдаваемый текст часто содержит фактические ошибки. На середину 2022 года технологии генеративного ИИ, если обратиться к кривой инноваций Gartner, находятся на пике раздутых ожиданий, а не приближаются к плато продуктивности.

### Цикл хайпа в сфере ИИ



### Риски и проблемы

Помимо постоянно растущих требований к вычислительным мощностям, существуют еще несколько категорий рисков, которые могут затормозить развитие рынка генеративных ИИ.

Во-первых, в среднесрочной перспективе будет сложно отделить тексты, составленные людьми от сгенерированных машинами. И эти данные будут попадать в обучающую выборку самих нейросетей. Это может привести к тому, что модель прекратит извлекать новое из поступающей информации и будет лишь укреплять уже сложившиеся паттерны, которые могут и не иметь никакого отношения к реальности. Иными словами, “буксовать” на месте.

Вторая проблема — юридические сложности, связанные с использованием данных, собранных в интернете, для обучения систем ИИ. Компания Getty Images в начале 2023 года подала иск против разработчика системы генерации изображений Stable Diffusion, в котором она обвиняет создателей в нарушении авторских прав на свои изображения. Stable Diffusion, наряду с другими изображениями, использовала 12 млн фотографий из базы данных Getty Images, не получив на это разрешения и не заплатив компенсацию.

Медиа-компании CNN и Wall Street Journal подали схожий иск к OpenAI, так как она использовала тексты статей этих и других издательств для обучения своих моделей.

Также подано несколько исков от программистов, хранящих исходный код своих программ на GitHub. Система Copilot от Microsoft, которая подсказывает написание кода, предлагает пользователям решения, полностью скопированные из чужих программ, даже если их лицензия не разрешает такое использование.

Тексты и изображения, которые созданы с помощью ИИ, не подпадают под защиту авторского права, считает Бюро по авторским правам США. Следовательно, корпорации в США могут столкнуться с неправомерным использованием своих материалов, которые будут “скармливаться” генеративному ИИ.

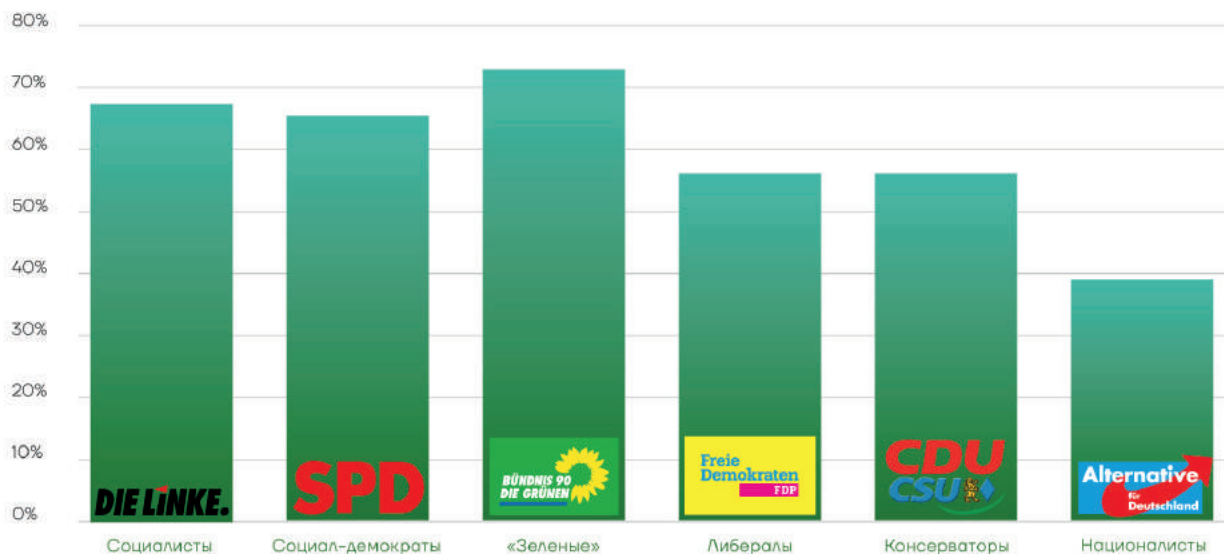
В ЕС развитие ИИ-систем также тормозится юридической неопределенностью. Европейская комиссия разрабатывает закон об ИИ с 2020 года. Но так как технологический ландшафт в этой области меняется слишком быстро, законодатели пока не могут сформулировать четких долгосрочных правил. И принятие этого закона постоянно откладывается.

Третья категория рисков связана с обработкой данных, относящихся к банковской, коммерческой, медицинской и другим видам тайн. Создание генеративных ИИ для финансовой или медицинской сфер требует доступа к обучающим данным, которые охраняются законом. Использование ChatGPT и подобных систем на середину марта 2023 года уже запрещено в JPMorgan, Citigroup, Wells Fargo и в программном обеспечении ЦБ Ирландии, чтобы не допустить утечек.

В организациях, работающих с чувствительной информацией, применение генеративных моделей ИИ возможно только в тех случаях, когда их обучение и использование происходит в рамках самого предприятия, а не отдается внешнему подрядчику. Однако для одной компании создание генеративной модели ИИ, полезной в практическом смысле, является слишком дорогой задачей.

Четвертая категория рисков возникает в политической области. Использование результатов работы генеративных ИИ может быть законодательно ограничено из-за того, что тексты не будут политически нейтральными, либо из-за несоответствия официальной позиции властей отдельно взятого государства. Исследователи из Мюнхенского технологического университета и Гамбургского университета Йохан Хартманн (Jochen Hartmann), Яспер Швенцов (Jasper Schwenzow) и Максимилиан Витте (Maximilian Witte) задали модели ChatGPT 630 политических вопросов, и ответы на них больше всего совпали с повесткой партии «Зеленых» и союза левых сил «Der Linke».

## Степень совпадения политической позиции ChatGPT с позициями политических партий Германии



Источник данных: The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation

## Реальные перспективы

Развитие ИИ, включая генеративные модели, сможет повысить производительность труда в тех областях, где цена ошибки небольшая. Генеративные модели уже находят широкое применение в маркетинге, копирайтинге, создании иллюстраций для блогов. По мере их дальнейшего развития и повышения качества создаваемых текстов, генеративные модели будут использоваться для написания черновиков-заготовок для компьютерных программ, юридических документов и других структурированных текстов. Однако их применение в других областях будет сдерживаться юридическими проблемами и сложностями с получением качественных обучающих данных.

На уровне государственного регулирования пока нет единого подхода. КНР сдерживает проникновение ChatGPT. WeChat в феврале 2023 года начал блокировать аккаунты, которые используют API чат-бота. По данным Nikkei Asia, Tencent и Ant Group получили распоряжение не предоставлять ни прямого, ни косвенного доступа к ChatGPT.

Наиболее выгодным для России видится создание собственных конкурентоспособных решений, а не запрет на использование зарубежных. Отечественные компании уже активно разрабатывают и внедряют генеративные нейросети, и если они привлекут клиентов, уставших от политической ангажированности западных решений, Россия окажется одним из лидеров не только на внутреннем, но и на глобальном рынке.