



УРАЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ
ЭКОНОМИЧЕСКИЙ
УНИВЕРСИТЕТ



КАФЕДРА
ШАХМАТНОГО
ИСКУССТВА
И КОМПЬЮТЕРНОЙ
МАТЕМАТИКИ

<https://vikchas.ru>

ХІХ сессия «Синтез права и технологий: на пути к технологическому лидерству»
4-5 июня 2026 года

**«Основы будущего: фундаментальные проблемы
генеративного ИИ»**

**«Правовые границы интеллекта регулирования
генеративного ИИ в системе обеспечения национальной
безопасности России»**

Часовских Виктор Петрович

Профессор, д.т.н., Академик АВН и РАН

Екатеринбург 2026

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Историю возникновения и развития искусственного интеллекта (ИИ) рассмотрим с начала исследования головного мозга человека.

Впервые нервная клетка была обнаружена в 1824 году французским биологом Рене Дютроше, термин «**нейрон**» для обозначения нервной клетки ввёл немецкий анатом Генрих Вальдейер в 1891 году.

В своей работе он описал нервную систему как совокупность отдельных, анатомически и генетически **независимых единиц** — **нейронов**.

В 1957 году Фрэнк Розенблатт, нейрофизиолог, разработал **перцептрон** — первую компьютерную модель нейрона головного мозга в Корнеллском аэрокосмическом институте и там же был собран первый в мире **нейрокомпьютер** «Марк-1». Демонстрация состоялась 23 июня 1960 года.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

«Марк-1» стал аппаратной реализацией перцептрона — одной из первых моделей **искусственных нейронных сетей**.

Перцептрон был первоначально программно смоделирован на компьютере **IBM 704** в **1957** году.

Развитие идей перцептрона привело к созданию:

- многослойного перцептрона (MLP);

- сверточных сетей (CNN);

- рекуррентных сетей (RNN, LSTM, GRU);

- трансформеров GPT, BERT и др.;

- генеративных моделей (GAN, VAE).

Понимание работы перцептрона — обязательный этап для любого специалиста в области ИИ в университетах Мира.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Искусственный интеллект в 50-е годы 20-го века развивался как область программного обеспечения для цифровых ЭВМ.

ЭВМ были основным инструментом для реализации идей моделирования интеллекта, а направленность на программирование определяла практическую значимость исследований.

Термин ИИ был предложен Джоном Маккарти в 1956 году при организации Дартмутского семинара.

Он написал в заявке на проведение семинара: «Мы предлагаем исследование искусственного интеллекта. Исследование основано на предположении, что любое другое интеллекта может в принципе быть столь точно описано, что машина сможет его **симулировать**.»

Джон Маккарти сосредоточился на **программном обеспечении** — он создал **LISP** как язык, подходящий для задач ИИ.

Участники Дартмутского семинара 1956 года — ключевого события в истории искусственного интеллекта



Развитие ИИ

Ранний этап (1940–1950-е)

1943: Уоррен Маккалок и Уолтер Питтс предложили модель искусственного нейрона.

1950: Алан Тьюринг опубликовал статью “Может ли машина мыслить?” и **предложил игру в имитацию.**

1956: Конференция в Дартмуте, где Джон Маккарти впервые использовал термин “искусственный интеллект”.

Период оптимизма и первые достижения (1950–1970-е)

1950–1960-е: Первые программы, такие как **Логик-теоретик** (доказательство теорем) и **ELIZA** (чат-бот).

1960-е: Попытки создания универсального ИИ, но ограниченные мощности и данные стали препятствием.

Первая зима ИИ (1970-е) ❄️

Причины: Провал проектов (например, машинного перевода), ограниченные возможности алгоритмов.

Итог: сокращение финансирования, скептицизм.

Возрождение: экспертные системы и машинное обучение (1980–1990-е)

1980-е: Экспертные системы (например, **MYCIN** для диагностики) стали коммерчески успешными.

1980–1990-е: развитие нейронных сетей и машинного обучения, но технологии ещё не были достаточно мощными.

Вторая зима ИИ (конец 1980-х – 1990-е) ❄️

Причины: Экспертные системы оказались дорогими и нефлексибельными, нейронные сети не оправдали ожиданий.

Итог: снова сокращение финансирования и интереса.

21 век: глубокое обучение и Big Data (2000-е – настоящее время) и ГИИ

2000-е: Прорыв в глубоком обучении благодаря GPU и большим данным.

2012: AlexNet побеждает в ImageNet, начинается бум свёрточных сетей (CNN).

2017: Трансформеры (Transformer) — основа GPT и BERT.

2020-е: ChatGPT, DALL·E, мультимодальные модели, автономные системы.

Третья зима ИИ? (Инвестиции 0,1 млрд. 2003 г. 2025 – 350 млрд.) ❄️ ?

Возможные причины:

- ✓ Ограничения больших моделей (галлюцинации, высокая стоимость обучения).
- ✓ Этические и регуляторные вопросы (права, приватность, авторство).
- ✓ Экономические факторы (снижение инвестиций после бума).

Будет ли? Пока ГИИ активно развивается, но возможен спад, если не будет новых прорывов.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ (ИИ) — это СПЕЦИАЛЬНОСТЬ ГРУППЫ СПЕЦИАЛЬНОСТИ «КОМПЬЮТЕРНЫЕ НАУКИ И ИНФОРМАТИКА (COMPUTER SCIENCE)» ОБЛАСТИ НАУКИ «ЕСТЕСТВЕННЫЕ НАУКИ», разрабатывающей методы, программные приложения и вычислительные системы, которые способны воспроизводить и имитировать человеческое поведение (оценивает тест Тьюринга).

Основные характеристики ИИ включают следующее:

Анализ

Обучение

Действие

Слабый и сильный ИИ

В 1980 году философом Джоном Сёрлем была опубликована в журнале «Behavioral and Brain Sciences» статья “**Minds, Brains, and Programs**”, в которой были определены понятия **слабый и сильный ИИ**.

Эти термины стали частью философской дискуссии о природе искусственного интеллекта и были приняты в практику ИИ.

Сёрль использовал эти понятия в рамках “Китайской комнаты”— мысленного эксперимента, критикующего утверждение, что компьютерная программа может обладать сознанием.

Слабый ИИ – это системы, которые лишь имитируют интеллектуальное поведение, но не обладают сознанием (например, чат-боты, алгоритмы распознавания образов).

Сильный ИИ – гипотетический ИИ, который действительно понимает и осознаёт себя как человек (аналог общего искусственного интеллекта, AGI).

Сегодня **слабый ИИ**– это практически все существующие системы (от GPT до беспилотных авто), а **сильный ИИ** остаётся теоретической концепцией.

Сильный, слабый и генеративный ИИ

Что общее?

Цель: все три типа ИИ направлены на автоматизацию задач и улучшение функциональности.

Технологическая основа: используют схожие методы и алгоритмы (например, нейронные сети, логический вывод, обработка данных).

Развитие: эволюция слабого ИИ и генеративного ИИ может приблизиться к сильному ИИ в будущем.

Что разделяет?

Слабый ИИ: решает только узкие задачи (например, распознавание изображений).

Генеративный ИИ: решает комплексные задачи, объединяя узкие функции (например, робот с компьютерным зрением, NLP и планированием).

Сильный ИИ: способен обобщать знания и решать любые задачи, как человек.

Генеративный искусственный интеллект

Точная дата появления ГИИ не зафиксирована в истории, так как его развитие было постепенным и связано с эволюцией отдельных технологий.

Однако можно говорить о 1980–1990-х годах как о периоде, когда начали активно развиваться нейронные сети, экспертные системы и другие технологии, которые позже стали компонентами генеративных систем.

В эти годы стало ясно, что классические подходы к созданию интеллектуальных вычислительных машин исчерпали себя и потребовались новые решения с функцией самообучения.

Генеративный искусственный интеллект

Ярким примером ГИИ служат такие большие языковые модели (LLM, Large Language Models), как **T5** (*преобразователь передачи текста в текст*), **GPT** (*генеративный предобученный трансформер*) и **Llama** (*семейство больших языковых моделей с открытым исходным кодом*), специализирующиеся на генерации и понимании текста.

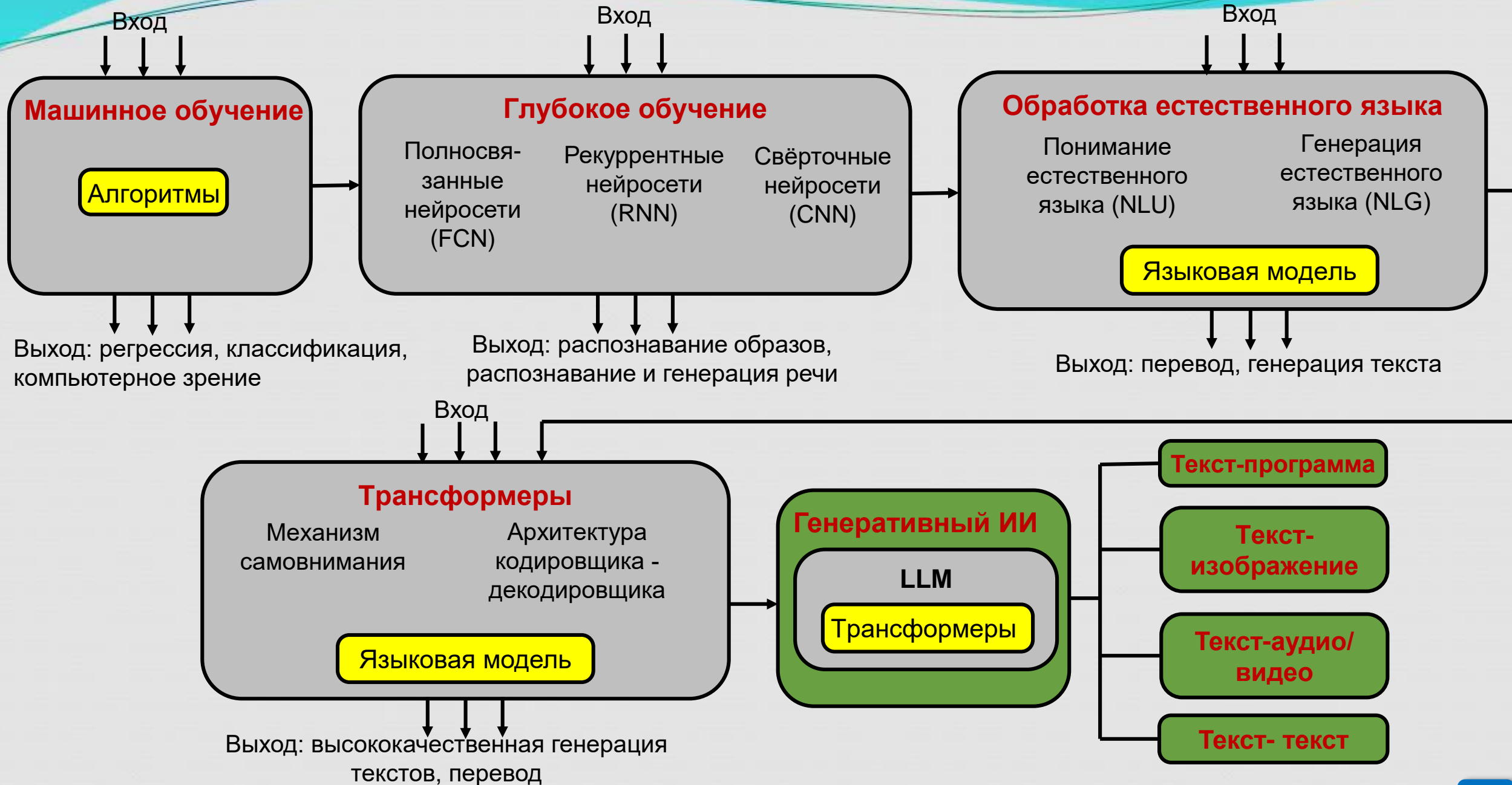
На основе методов и механизмов глубокого обучения такие модели выявляют закономерности и представления признаков из огромного количества существующих данных. ГИИ часто рассматриваются как подмножество ИИ-систем глубокого обучения.

Генеративный искусственный интеллект

Генеративный искусственный интеллект (ГИИ, Generative AI, GenAI) — направление ИИ, сформировавшееся благодаря многолетним исследованиям в смежных ИТ-дисциплинах и ориентированное на **автоматическое создание новых данных, контента и знаний**.

Он объединяет потенциал машинного и глубокого обучения, а также инструменты для мультимедийной генерации креативного контента.

В отличие от аналитических моделей, ограниченных выявлением паттернов и прогнозированием, генеративный искусственный интеллект способен самостоятельно продуцировать принципиально новый контент — **тексты, графику, аудио, видео и программный код**, комбинируя и трансформируя сведения, извлечённые из обучающих наборов данных.



Машинное обучение

Машинное обучение — это ключевая технология искусственного интеллекта, позволяющая компьютерам “учиться” на основе данных без явного программирования.

Основные типы машинного обучения - обучение с учителем; обучение без учителя; обучение с подкреплением.

Ключевые технологии - нейронные сети (глубокое обучение, трансформеры)

Ансамблевые методы - случайный лес, градиентный бустинг (*новые модели из более мелких на основе градиента функции потерь*)

Байесовские методы - вероятностные модели.

Машинное обучение существенно меняет многие отрасли — от медицины до финансов, и продолжает стремительно развиваться.

Глубокое обучение

Глубокое обучение (DL) — это подраздел машинного обучения, основанный на использовании многослойных нейронных сетей для решения сложных задач.

Для глубокого обучения используют искусственные нейронные сети, с помощью которых выявляются закономерности в данных и связи между данными:

- ✓ полносвязанные нейросети (FCN);
- ✓ сверточные сети (CNN) - для работы с изображениями;
- ✓ рекуррентные сети (RNN, LSTM, GRU) - для последовательных данных;
- ✓ Трансформеры - для обработки текста и других последовательностей.

Ключевое различие заключается в том, что алгоритмы машинного обучения учатся на данных, выявляя простые закономерности и взаимосвязи, в то время как для глубокого обучения применяются искусственные нейронные сети с несколькими слоями (отсюда и слово «глубокое» в названии), взаимосвязанных узлов для изучения сложных закономерностей и представлений (признаков).

Такие слои позволяют алгоритму постепенно все глубже анализировать данные и формировать некое понимание.

Обработка естественного языка

Обработка естественного языка (NLP) — одно из направлений ИИ, посвященное компьютерному **анализу** (пониманию и интерпретации) и **синтезу** (генерации) человеческих высказываний и осмысленных текстов на естественных языках. Сферы NLP и глубокого обучения тесно связаны между собой, причем модели глубокого обучения представляют собой мощные инструменты, значительно расширяющие возможности NLP. Взаимосвязь их следующая.

Нейронные сети. Для глубокого обучения используются нейронные сети с несколькими слоями, моделирующие сложные закономерности в данных.

Рекуррентные нейронные сети (RNN) обрабатывают последовательные данные и выявляют зависимости в тексте.

Сверточные нейронные сети (CNN). CNN использовались для обработки изображений, но их стали применять и для текста, так как сверточные сети позволяют хорошо выявлять локальные закономерности в тексте.

Трансформеры

Модель трансформера для выявления сложных связей между словами в предложении использует механизмы самовнимания.

Основные компоненты - это:

Механизм самовнимания - позволяет модели оценивать важность различных слов в предложении относительно друг друга и выявлять сложные взаимосвязи.

Архитектура кодировщика-декодировщика - необходима для таких задач, как машинный перевод.

Известные модели на основе трансформеров:

- ✓ **BERT**: двунаправленный трансформер-энкодер (преобразовывает входную последовательность данных, например, текст в числовое представление) для понимания языка;
- ✓ **GPT(1-5)**: однонаправленный трансформер-декодер для генерации текста;
- ✓ **T5**: трансформер, представляющий все NLP-задачи как задачи “текст-в-текст”;
- ✓ **DALL-E, Stable Diffusion**: генерация изображений по текстовому описанию;
- ✓ **ViT**: трансформеры для компьютерного зрения.

Генеративный искусственный интеллект

В 2017 году, благодаря разработке трансформеров в области обработки естественного языка, был совершен значительный прорыв. Новая схема с механизмом самовнимания позволила моделям эффективнее обрабатывать большие объемы текста с фокусировкой на наиболее значимых его частях.

В результате трансформеры стали основой, на которой были созданы многие современные генеративные модели ИИ.

С появлением огромных массивов данных и увеличением вычислительной мощности исследователи стали обучать генеративные модели на огромных объемах данных.

Такие большие языковые модели, как GPT от OpenAI и BERT от Google, обучаются на огромных текстовых наборах, благодаря чему генерируют осмысленный и связный текст или вступают в диалог с пользователями, как это происходит в чат-ботах, работающих на основе ИИ.

Генеративный искусственный интеллект

ГИИ обладает огромным количеством преимуществ, а сферы его применения обширны и разнообразны.

Наиболее значимые преимущества.

Творчество и инновации. ГИИ помогает создавать креативный контент — такой как художественные, музыкальные и литературные произведения. Он помогает художникам, музыкантам и писателям исследовать новые идеи и облегчает творческий процесс.

Повышение эффективности и автоматизация. Генеративный ИИ автоматизирует рутинные задачи и процессы, что позволяет экономить время и сокращать количество человеческих ошибок. Например, он может составлять отчеты, создавать маркетинговые материалы и обрабатывать запросы в службу поддержки клиентов.

Генеративный искусственный интеллект

Персонализация. ГИИ может адаптироваться к предпочтениям конкретных пользователей, повышая тем самым их удовлетворенность от общения с ним. Так, например, он может выдавать персональные рекомендации на потоковых сервисах, показывать персонализированную рекламу и составлять планы индивидуального обучения.

Создание контента. ГИИ позволяет быстро и в больших объемах создавать качественный контент. Например, с его помощью можно генерировать реалистичные изображения, писать статьи или создавать видео, что очень важно для отраслей, ориентированных на контент.

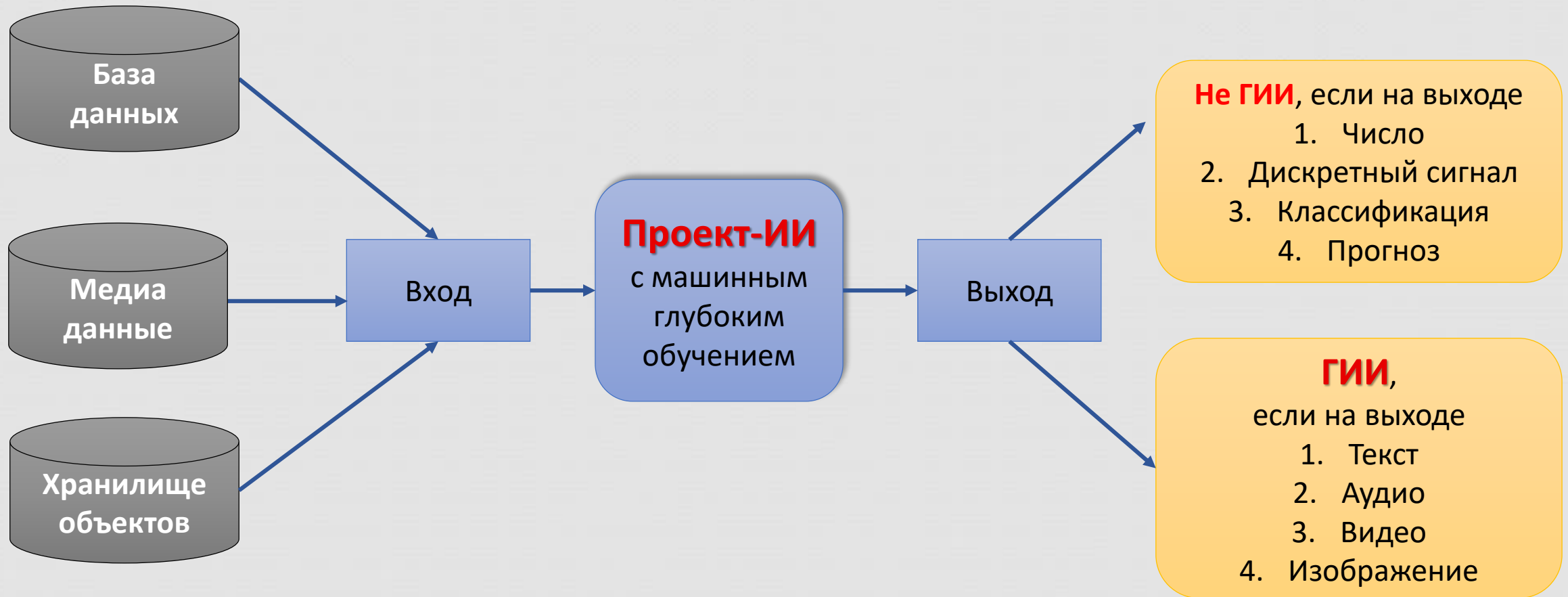
Генеративный искусственный интеллект

Аугментация (увеличение) данных. ГИИ может создавать синтетические данные для дополнения реальных наборов данных, что бывает полезно для настройки моделей машинного обучения, особенно в тех случаях, когда реальных данных мало или они конфиденциальны.

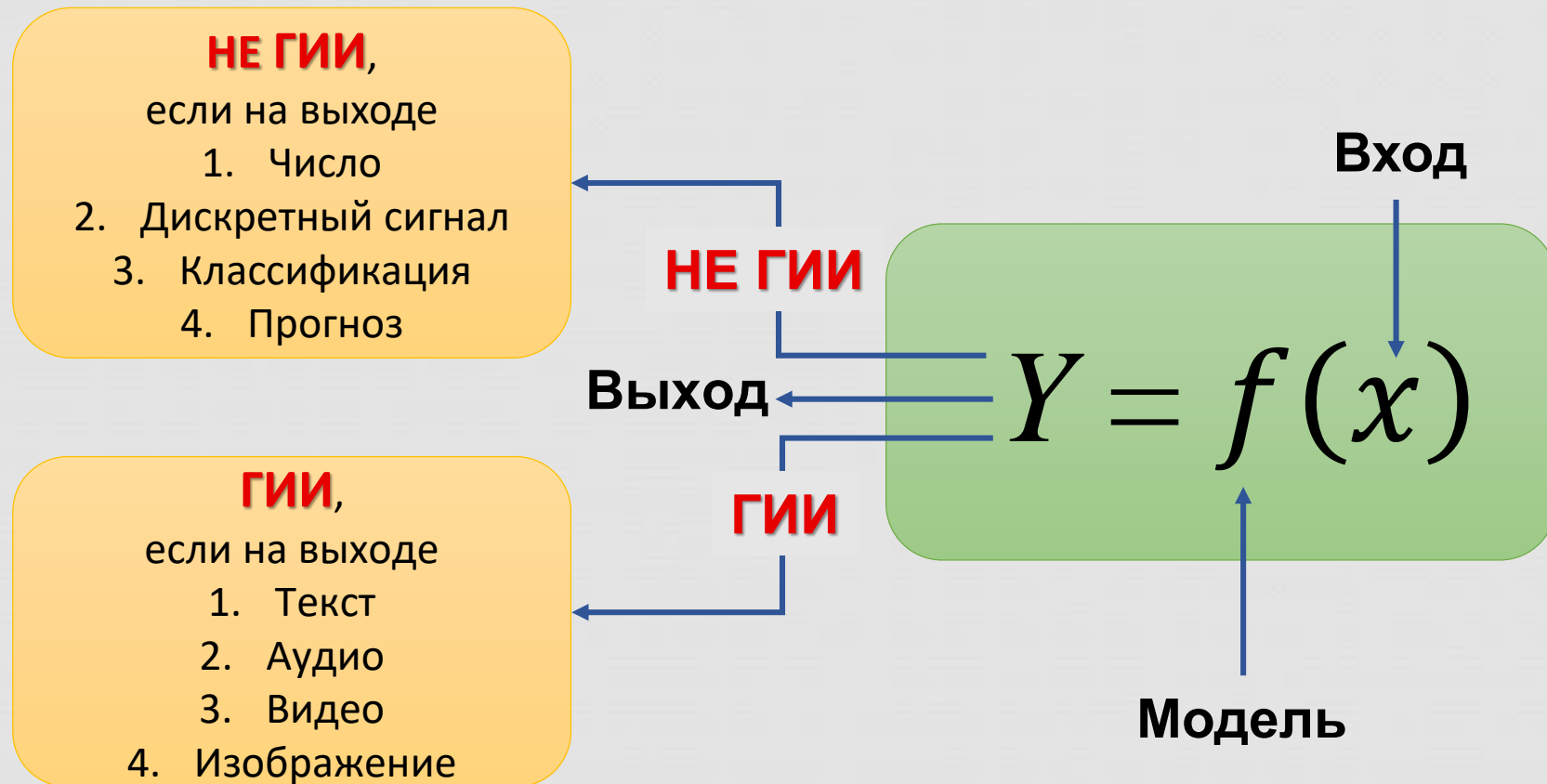
Повышение доступности. С помощью генеративного ИИ можно создавать инструменты и приложения, улучшающие доступность продукции и услуг — например, генерировать текстовые описания для изображений или озвучивать видео для людей с нарушениями зрения.

Исследование и моделирование. Генеративные модели могут имитировать различные сценарии и среды, что бывает полезно для проведения научных исследований, виртуального обучения и планирования.

Что относится к ГИИ?



Математическое представление



Генеративный искусственный интеллект

Уравнение $Y = f(x)$ символически означает некий способ получения результата (Y) на основе различных данных (x).

В данном случае Y — это результат работы ИИ-системы («выход»), f — это некая функция или модель, используемая для расчетов, а x может означать все, что угодно — от необработанного текста (например, статьи в журнале) до CSV-файлов (например, годовых отчетов) или даже изображений.

Выход модели определяется всеми входными данными. Если Y представляет собой число — например, прогнозируемую цену автомобиля, то данная модель не считается генеративной.

Однако если Y представляет собой предложение — например, «Карта депонированного углерода»), то модель уже можно причислить к генеративным.

Категории генеративного ИИ

С 2024 года наблюдается широкий спектр моделей генеративного ИИ, включая следующие:

1. Модели «текст-текст». Такие модели принимают текст на вход и генерируют текст на выходе. Тип текста может варьироваться — от статей на естественном языке и стихов до программного кода и даже документов HTML. Пример таких нейросетей — **Gemini** от Google, **GPT-4**, **Claude Opus** и **LLaMA 3.1**.

2. Модели преобразования «текст-изображение». Эти модели генерируют изображения на основе текстового описания (промпта). Например, можно предложить описание ели или кедра с красными ветками — и будет сгенерировано соответствующее изображение. Пример такой нейросети — **Midjourney**.

3. Модели «изображение-изображение». Эти модели на основе предоставленного изображения выдают его измененную версию или комбинируют изображения.

Категории генеративного ИИ

4. Модели «изображение-текст». Эти модели генерируют текст на основе изображений, что бывает особенно полезно для таких задач, как перевод презентаций в текстовый вид. Пример модели такого типа — LLaVA.

5. Модели «речь-текст». Эти модели транскрибируют речь в текст, что бывает полезно для таких задач, как запись совещаний — например, при общении в Zoom. В качестве примера можно привести Whisper и DeepSpeech.

6. Модели преобразования «текст-аудио». Эти модели на основе текстовых промптов генерируют музыку или звуки. Особенно они полезны для аудиотранскрипции. Пример модели такого типа — **Aura** от Deepgram.

7. Модели «текст-видео». Такие модели на основе текстовых промптов генерируют видео. В ближайшем будущем мы, возможно, увидим целые фильмы, созданные на основе текстовых сценариев. Примеры таких моделей — **Sora, Dream Machine** и **Kling**.

Современная классификация генеративного ИИ

1. Фундаментальные модели — это большие, предварительно обученные модели *машинного обучения*, служащие основой для различных ИИ-приложений. Такие модели обучаются на обширных и разнообразных наборах данных (текст, изображения или аудио), и их можно настраивать для решения конкретных задач с минимальным дополнительным обучением. Благодаря своему масштабу и универсальности, фундаментальные модели адаптируются для решения широкого спектра задач — от перевода с одного языка на другой (*но машинный перевод не улавливает нюансы, стилистические особенности, культурный контекст и эмоциональную окраску текста*) до распознавания изображений.

В качестве примера можно привести **DALL-E** и **Gato** от DeepMind — это фундаментальные модели, обученные на изображениях и текстах и генерирующие оригинальные изображения на основе текстовых описаний.

Современная классификация генеративного ИИ

2. Большие языковые модели (LLM) — это разновидность фундаментальных моделей, которые используют механизмы глубокого обучения для анализа огромных объемов данных и обучения на них. В результате LLM достигают исключительной эффективности в обработке и генерации текста почти неотличимого от составленного человеком. По сути, основываясь на распознанных на основе полученных данных закономерностях и структурах, они способны создавать совершенно новые текстовые комбинации, отражающие нюансы и сложности естественного языка.

Лучшие примеры **LLM** — **GPT (1-5), LLaMA, DEEPSEEK v.5**

Ключевые характеристики LLM

1. Разработаны специально для решения задач, связанных с естественным языком.
2. Генерируют текст, обрабатывают язык и понимают языковые шаблоны.
3. Обучаются на массивных наборах данных из различных источников.

LLM — это особый тип фундаментальной модели. Как фундаментальные модели так и LLM представляют собой ведущие системы генеративного ИИ, благодаря которым ЭВМ создают разнообразный инновационный контент.

В настоящее время термин «**фундаментальная модель**» часто используется в значении «**большая языковая модель**», поскольку языковые модели — это самый яркий пример систем с широкими возможностями, которые можно адаптировать для различных целей.

Ключевое различие между ними заключается в том, что термин «**большие языковые модели**» относится к системам, ориентированным именно на решение языковых задач, в то время как термин «**фундаментальная модель**» имеет более широкое значение и означает систему, на основе которой в будущем могут появиться новые типы систем.

ИИ-агенты

Под широким термином «**ИИ-агенты**» подразумеваются автономные или полуавтономные системы, предназначенные для выполнения конкретных задач или решения проблем посредством анализа окружения, обработки информации, принятия решений и выполнения определенных действий.

В контексте генеративного ИИ и LLM **ИИ-агенты** — это сложные системы, использующие генеративные модели, такие как LLaMA, Mistral и GPT-4 для выполнения сложных задач, связанных с обработкой естественного языка (его пониманием, генерацией и взаимодействием с пользователем).

Эти агенты предназначены для имитации взаимодействия с человеком, автоматизации процессов принятия решений и облегчения работы различных приложений — от создания контента до персонализации пользовательского опыта.

ИИ-агенты

Основные цели ИИ-агентов ИИ в области генеративного ИИ следующие.

1. Улучшение взаимодействия с пользователями. ИИ-агенты могут поддерживать диалог и выступать в роли виртуальных помощников, обеспечивая тем самым более естественное и увлекательное взаимодействие с пользователями. Например, ИИ-ассистент, настроенный на медицинские данные, может давать советы по вопросам здравоохранения.

2. Автоматизация сложных процессов. Эти агенты могут автоматизировать сложные рабочие процессы, понимая инструкции пользователя, последовательно выполняя задания и сообщая о результатах. В качестве примера можно представить ИИ-инструмент на основе промпта, генерирующий рекламный текст, публикации в блогах или контент для социальных сетей.

ИИ-агенты

3. Динамическая интеграция знаний. ИИ-агенты могут получать доступ к внешним источникам данных в режиме реального времени, что позволяет им предоставлять актуальную и точную информацию, выходящую за рамки статических знаний, закодированных в их обучающих данных. Пример — агент-фреймворк, в котором разные модули отдельно обрабатывают генерацию текста, вызовы API и управление пользовательским интерфейсом.

4. Персонализация. ИИ-агенты могут со временем изучать предпочтения и поведение пользователей, чтобы предлагать персонализированные рекомендации или предоставлять помощь с учетом индивидуальных потребностей. Пример — бот для обслуживания клиентов, улучшающий свои ответы на основе взаимодействия с пользователем и отзывов.

ИИ-агенты

5. Поддержка принятия решений и решение проблем. ИИ-агенты могут анализировать большие объемы данных, выявлять закономерности и предоставлять выводы, рекомендации или резюме, облегчая тем самым процесс принятия решений.

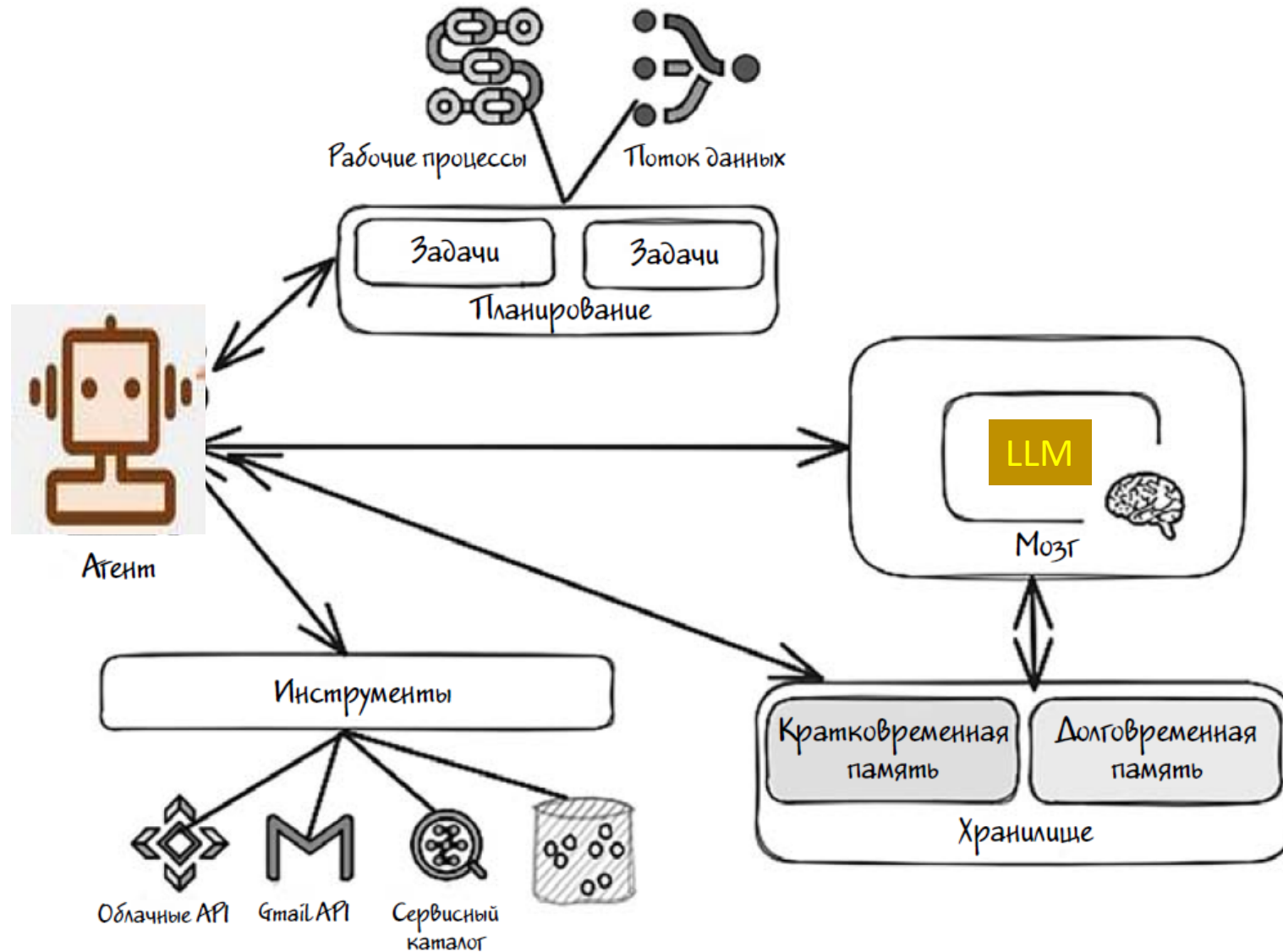
Пример — агент, отвечающий на технические запросы, извлекающий данные из научных баз данных и излагающий информацию на естественном языке.

ИИ-агенты

Архитектура LLM-агента состоит из четырех основных компонентов:

1. **Мозг**. Ядро агента, обычно LLM, действующая как центр принятия решений.
2. **Планировщик**. Этот компонент разбивает сложные задачи на управляемые шаги.
3. **Память**. Агент сохраняет контекст и информацию для последующего использования, помогая поддерживать непрерывность.
4. **Инструменты**. Агент использует различные инструменты, осуществляющие вызовы API, запросы к базе данных и даже создающие программный код для выполнения указанных пользователем задач.

ИИ-агенты



Разработка ИИ-агентов

Разработка ИИ-агентов — одно из передовых направлений в области искусственного интеллекта. ИИ-агенты — это комплексные системы, использующие генеративные модели для выполнения сложных задач, связанных с обработкой естественного языка и со взаимодействием с пользователем на естественном языке.

Они создаются для имитации взаимодействия с человеком, для автоматизации процесса принятия решений и для выполнения широкого спектра прикладных задач — от создания контента до персонализации пользовательского опыта.

По мере того, как бизнес-приложения становятся все более зависимыми от взаимосвязи различных систем, ИИ-агенты смогут пользоваться возможностями LLM для автономного или полуавтономного выполнения задач.

Такие агенты оптимизируют рабочие процессы благодаря способностям языковых моделей понимать и генерировать текст на естественном языке и их сочетанию с внешними инструментами или средами, предоставляющими возможности дополнительной памяти, планирования и взаимодействия с другими инструментами.

Ключевые этапы и направления развития ИИ-агентов ГИИ

На основе существующих трендов и прогнозов можно выделить

Интервал развития	Результаты ИИ-агенты ГИИ
2015–2020 годы: зарождение и первые шаги	ИИ-агенты находились в стадии экспериментальных разработок. Основные усилия были сосредоточены на улучшении машинного обучения, обработке естественного языка (NLP) и базовых моделей. Появились первые системы, способные выполнять простые задачи по заданным сценариям.
2021–2025 годы: прорыв и активное внедрение	Развитие архитектуры Transformer (с 2017 года) радикально изменило подход к обработке естественного языка. В 2018 году появились первые модели на её базе — GPT-1 и BERT. Рост интереса к генеративному ИИ (ГИИ). В 2022 году OpenAI запустила публичную версию ChatGPT на основе GPT-3.5. Активное внедрение в бизнес: к 2025 году ИИ-агенты начали активно использоваться в корпоративных процессах.
2026–2030 годы: масштабирование и мультиагентные системы	Рост автономности: ИИ-агенты перейдут от роли «помощника» к статусу «исполнителя» , самостоятельно планирующего и выполняющего многошаговые задачи. Применение в новых сферах: к 2030 году ожидается, что агенты будут управлять производственными процессами, анализировать инвестиции, оптимизировать бюджеты, а также использоваться в исследованиях, разработке продуктов и маркетинге.
2031–2035 годы: глубокая интеграция и новые горизонты	Субъекты экономики: ИИ-агенты могут стать полноценными участниками экономических процессов — автономными игроками на стороне производства и потребления. Прогнозируется, что ИИ-агенты будут дополнять человеческие способности, освобождая людей для более творческой и стратегической работы.

Развитие ИИ-агентов

1. Агенты, выполняющие конкретные задачи. ИИ-агенты первой волны будут выполнять специализированные, четко сформулированные задачи. Такие агенты предназначены для выполнения повторяющихся или рутинных процессов, и благодаря им повысится эффективность в конкретных областях применения. Так, например, ИИ-агент может обрабатывать запросы клиентов в колл-центре, автоматизировать ввод данных или генерировать отчеты на основе структурированных данных. Эти первые ИИ-агенты будут отлично справляться с одним заданием, что позволит предприятиям автоматизировать мелкие, изолированные задачи.

2. ИИ-агенты помощники. Следующим шагом станет разработка более универсальных ИИ-помощников, способных работать и повышать производительность вместе с человеком. Такие агенты будут не просто выполнять задания самостоятельно, но и активно сотрудничать с людьми, внося свои предложения, предоставляя информацию в режиме реального времени и автоматизируя часть более сложных задач. Например, ИИ-помощники смогут помогать маркетологу составлять материалы для рекламной кампании, а разработчику программного обеспечения — писать код, что ускорит выполнение работы.

Развитие ИИ-агентов

3. Многоагентные системы. На самом продвинутом этапе мы станем свидетелями появления многоагентных систем. Они будут состоять из нескольких ИИ-агентов, работающих совместно для поддержания сложных последовательных и взаимосвязанных процессов, охватывающих различные отделы или рабочие циклы. На этой стадии ИИ-агенты будут координировать свои действия между собой и выполнять такие масштабные операции как управление всей цепочкой поставок или многоступенчатым процессом утверждения. Агенты смогут автономно делегировать друг другу задачи и адаптироваться к специфическим требованиям различных этапов рабочего цикла.

Фреймворки для разработки ИИ-агентов

Список фреймворков для разработки ИИ-агентов можно найти по ссылке: <https://github.com/e2b-dev/awesome-ai-agents> .

Многие из указанных инструментов специализированы для решения конкретных задач и не требуют предварительных знаний программирования, в то время как некоторые предназначены для решения более широкого круга задач и требуют развитых навыков программирования.

Наиболее распространенные фреймворки с открытым исходным кодом для создания ИИ-агентов на основе больших языковых моделей (LLM).

На сайте <https://vikchas.ru> приведены 2 примера создания ИИ-агента для построения карт депонирования углерода лесом с использованием фреймворка LangChain и LLM.

Фреймворки для разработки ИИ-агентов

1. Langchain и LangGraph

LangChain — это *полный* фреймворк, предназначенный для упрощения разработки приложений, подразумевающих интеграцию LLM с внешними источниками данных, рабочими процессами или API. Его основная цель — создать систему на основе цепочек, способную выполнять такие задачи, как ответы на вопросы, выполнение команд и взаимодействие с внешними системами с помощью моделей типа GPT или локальных моделей вроде LLaMA. Кроме того, в рамках проекта создан новый подпроект под названием LangGraph для разработки сложных агентных систем.

2. CrewAI

CrewAI — это еще один широко используемый высокоуровневый фреймворк, предназначенный для ролевых мультиагентных взаимодействий. Он ориентирован на создание ИИ-агентов, работающих над сложными задачами сообща, и управление ими. Фреймворк позволяет координировать работу сразу нескольких ИИ-моделей или ИИ-агентов, что делает его идеальным для задач, требующих разносторонней

Достоверность ИИ

Основным элементом построения проектов ИИ и формирования ответа пользователю является искусственный нейрон, его математическая модель. Нейроны современных проектов ИИ показаны в следующей таблице:

Модель нейрона	Год	Ключевая особенность
Маккаллок — Питтс	1943	Пороговая активация
Перцептрон	1957	Обучение с коррекцией ошибок
Многослойный перцептрон MLP	1986	Многослойность + нелинейность
Нейрон с различными функциями активации	1943	Способности адаптироваться и обучаться на основе различных входных данных
Радиально-базисный нейрон RBF	1988	Радиальные функции
Рекуррентные нейроны RNN	1990	Память состояний
Long Short-Term Memory LSTM	1997	Долговременная память
Нейроны в сверточных сетях CNN	1998	Локальная связность
Нейроны с механизмом внимания Transformer	2017	Механизм внимания

Достоверность нейронных сетей после машинного обучения

После завершения обучения нейронных сетей их достоверность (точность, надежность) зависит от нескольких ключевых факторов:

1. Качество обучающих данных:

- ✓ репрезентативность выборки;
- ✓ отсутствие или минимизация шума и выбросов;
- ✓ корректная разметка данных.

2. Архитектура модели:

- ✓ сложность модели должна соответствовать сложности решаемой задачи;
- ✓ правильно подобранные гиперпараметры.

3. Метрики эффективности:

- ✓ на тренировочном наборе: обычно 95-99% для хорошо обученных моделей;
- ✓ на валидационном наборе: обычно 85-95%;
- ✓ на **тестовом наборе: обычно 80-95%** для хорошо обобщающих моделей.

4. Проблемы обучения:

- ✓ переобучение (overfitting): модель “запоминает” тренировочные данные вместо обобщения;
- ✓ недообучение (underfitting): модель слишком проста для задачи.

Достоверность нейронных сетей при работе с реальными данными после обучения

Работа с реальными данными после обучения существенно отличается от процесса обучения:

1. Распределение данных:

- ✓ реальные данные могут отличаться от тренировочных (проблема смещения данных);
- ✓ достоверность снижается при работе с данными из “других распределений”.

2. Типичные показатели в реальных условиях:

- ✓ **как правило, 60-80% от достоверности на тестовом наборе;**
- ✓ может значительно снижаться со временем из-за изменения характеристик данных.

3. Факторы, влияющие на достоверность при инференсе (ранее обученную модель для принятия решений на основе новых данных.):

- ✓ доверительный уровень предсказаний;
- ✓ калибровка вероятностей;
- ✓ робастность к шуму и вариациям входных данных.

Оценка достоверности ГИИ по данным опросов пользователей в 2025 г.

Результаты опросов ВШЭ (Россия)

Общая достоверность: 70-75% пользователей отмечают приемлемый уровень достоверности ответов ГИИ.

Научная информация: достоверность оценивается в 65-70%.

Фактическая информация: около 60% респондентов сталкивались с фактическими ошибками.

Доверие к источнику: 55% пользователей проверяют информацию из ГИИ через дополнительные источники.

Результаты опросов ВВС (Великобритания)

Общественное восприятие: 62% опрошенных британцев считают ответы ГИИ «в целом достоверными».

Критическое отношение: 78% признают необходимость проверки информации.

Различие по возрастным группам: молодежь (18-24 года) оценивает достоверность выше (около 70%), люди старше 55 лет — ниже (около 45%).

Профессиональное использование: 55% специалистов используют ГИИ в работе, но только 40% полностью доверяют получаемой информации.

Благодарю за внимание!



Приложения

В Большой российской энциклопедии (БРЭ, издание 2004–2017 годов) информатика определяется как «наука об извлечении информации из сообщений, создании информационных ресурсов, программировании поведения машин, о других сущностях, связанных с построением и применением человеко-машинной среды решения задач моделирования, проектирования, взаимодействия, обучения и др.».

Научные специальности ВАК:

1. Естественные науки

1.2. Компьютерные науки и информатика – группа специальностей

1.2.1. Искусственный интеллект и машинное обучение - специальность

1.2.2. Математическое моделирование, численные методы и комплексы программ

1.2.3. Теоретическая информатика

1.2.4. Кибербезопасность

ИИ – это специальность группы специальности «Компьютерные науки и информатика», области науки «Естественные науки».

Указ Президента РФ от 10.10.2019 г. № 490

Искусственный интеллект - комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма), и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений;

Текущее состояние

На сегодняшний день **сильный ИИ** остаётся теоретической концепцией. Современные системы (включая продвинутые нейросети вроде GPT) относятся к слабому ИИ: они хорошо справляются с отдельными задачами, но не обладают универсальностью, самосознанием и способностью к подлинному обобщённому мышлению. Учёные предполагают, что создание AGI может занять десятки или даже сотни лет — либо оказаться принципиально недостижимым

ПРОЕКТ

ФЕДЕРАЛЬНЫЙ ЗАКОН

Об основах государственного регулирования сфер применения технологий искусственного интеллекта в Российской Федерации

(26 страниц)

Настоящий Федеральный закон вступает в силу с 1 сентября 2027.