МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
[Полное название вуза]
Факультет информационных технологий

РЕФЕРАТ

по дисциплине «Информатика»

на тему: «Искусственный интеллект: история развития и области применения»

Выполнил(а): студент гр. ____
_____ / И. О. Фамилия

Проверил(а): _____
_____ / И. О. Фамилия

[Город] 2026

Пример работы, выполненной в Универ AI · univerai.ru

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....3
1. История развития искусственного интеллекта.....5
1.1 Истоки и первые программы (1940–1970-е годы).....5
1.2 От экспертных систем к глубокому обучению.....8
2. Области применения искусственного интеллекта.....13
2.1 Искусственный интеллект в экономике, медицине и транспорте.....13
2.2 Образование, наука и государственное управление.....16
3. Принципы работы и границы возможностей искусственного интеллекта..19
3.1 Как обучается нейронная сеть: разбор простого примера.....19
3.2 Ограничения, риски и правовое регулирование.....21
ЗАКЛЮЧЕНИЕ.....26
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....28

ВВЕДЕНИЕ

Ещё десять лет назад разговор с компьютером на естественном языке казался приметой фантастического фильма. Сегодня студент просит чат-бота объяснить теорему, врач получает от программы подсказку на рентгеновском снимке, а банк одобряет кредит за несколько минут без участия сотрудника. По данным доклада AI Index Стэнфордского университета, доля организаций, использующих искусственный интеллект, выросла с 78% в 2024 году до 88% в 2025-м, а генеративный ИИ всего за три года достиг уровня распространения, на который персональному компьютеру и интернету потребовалось больше времени [6].

Стремительность этих перемен порождает и восторги, и тревогу. Одни видят в ИИ инструмент, который освободит людей от рутины, другие опасаются массовой потери рабочих мест, дипфейков и бесконтрольных решений машин. Чтобы трезво оценивать такие прогнозы, полезно знать, откуда взялась эта технология, через какие взлёты и провалы она прошла и где на самом деле приносит пользу.

Цель реферата — проследить историю развития искусственного интеллекта, охарактеризовать основные области его применения и показать, на чём основаны возможности современных систем и где проходят их границы. Для этого нужно решить четыре задачи: описать зарождение ИИ как научного направления и его первые достижения; рассмотреть переход от экспертных систем к машинному и глубокому обучению; проанализировать применение ИИ в экономике, медицине, транспорте, образовании, науке и государственном управлении; на простом числовом примере разобрать, как обучается нейронная сеть, и оценить риски технологии вместе с подходами к её правовому регулированию в России и Европейском союзе.

Объект исследования — искусственный интеллект как область информатики, предмет — этапы его развития, принципы работы и сферы практического использования. Работа состоит из введения, трёх глав, заключения и списка источников. В ней использованы классические статьи

А. Тьюринга и авторов Дартмутского проекта, учебники С. Рассела и П. Норвига и Л. Н. Ясницкого, материалы Большой российской энциклопедии, обзоры по глубокому обучению и распознаванию изображений, доклад AI Index 2026, нормативные акты Российской Федерации и Европейского союза.

1. История развития искусственного интеллекта

1.1 Истоки и первые программы (1940–1970-е годы)

Что такое искусственный интеллект? Большая российская энциклопедия определяет его как раздел информатики, который разрабатывает методы и средства решения на компьютере интеллектуальных задач, традиционно выполняемых человеком [2]. С. Рассел и П. Норвиг в своём известном учебнике показывают, что исследователи понимали эту цель по-разному: одни стремились к системам, которые думают или действуют как человек, другие — к системам, которые думают или действуют рационально [4]. От выбора подхода зависело, что считать успехом. Четыре трактовки, которые выделяют авторы, сведены на рисунке 1.



Рисунок 1 – Подходы к определению искусственного интеллекта

Источник: составлено автором по [4].

Тест Тьюринга, о котором речь пойдёт ниже, относится ко второй трактовке: машину оценивают по поведению, а не по устройству. Когнитивное моделирование, напротив, интересуется тем, совпадает ли ход рассуждений программы с ходом мысли человека, и потому тесно связано с психологией. Сами Рассел и Норвиг выбирают четвёртый вариант и описывают ИИ как науку о рациональных агентах — системах, которые

воспринимают среду и действуют так, чтобы добиться наилучшего ожидаемого результата [4].

Корни направления уходят в 1940-е годы. В 1943 году нейрофизиолог Уоррен Мак-Каллок и логик Уолтер Питтс предложили математическую модель нейрона, показав, что сеть таких элементов может вычислять логические функции. Почти одновременно появились первые электронные вычислительные машины, и вопрос о том, способны ли они на нечто большее, чем арифметика, перестал быть отвлечённым.

Самую известную постановку этого вопроса дал английский математик Алан Тьюринг. Его статья «Вычислительные машины и разум» 1950 года начинается словами: «Я предлагаю рассмотреть вопрос: могут ли машины мыслить?» [9]. Поскольку понятие «мыслить» трудно определить, Тьюринг заменил его игрой в имитацию, позже названной тестом Тьюринга: если человек-экзаменатор в переписке не может уверенно отличить машину от человека, машину разумно считать мыслящей. Учёный предсказал, что примерно через пятьдесят лет компьютеры будут играть в эту игру так хорошо, что средний экзаменатор после пяти минут вопросов будет иметь не более 70% шансов на верное опознание [9].

Рождением ИИ как научной дисциплины принято считать летний семинар 1956 года в Дартмутском колледже. Заявку на него 31 августа 1955 года подали Джон Маккарти, Марвин Минский, Натаниэль Рочестер и Клод Шеннон, и именно в этом документе впервые прозвучало словосочетание «искусственный интеллект» [8]. Авторы исходили из смелого предположения, что любую черту интеллекта в принципе можно описать настолько точно, чтобы машина могла её воспроизвести. Оптимизм был безграничным: участники рассчитывали добиться заметного прогресса за одно лето.

Первые результаты действительно впечатляли. В 1956 году А. Ньюэлл, Г. Саймон и К. Шоу представили программу «Логик-теоретик», доказывавшую теоремы из «Principia Mathematica» с помощью эвристик. В

конце 1950-х Маккарти создал язык программирования Лисп, на долгие десятилетия ставший основным языком исследователей ИИ. В 1957 году Фрэнк Розенблатт предложил перцептрон — математическую модель восприятия, а затем построил её электрическую реализацию [2].

Советская наука не стояла в стороне. В 1960-е годы В. Н. Вапник и А. Я. Червоненкис получили фундаментальные результаты о сходимости методов машинного обучения, которые позже легли в основу метода опорных векторов. Группа под руководством А. С. Кронрода в 1963 году начала работу над шахматной программой, которая в 1967 году выиграла у программы Маккарти телеграфный матч со счётом 3:1. Её наследница «Каисса» в 1974 году стала первым чемпионом мира среди шахматных программ [2].

Путь советских исследователей не был гладким. В начале 1950-х годов кибернетику в официальной печати называли буржуазной лженаукой, и лишь после 1955 года отношение к ней изменилось. В 1959 году при Президиуме Академии наук СССР был создан Научный совет по комплексной проблеме «Кибернетика», который возглавил академик А. И. Берг. В его рамках работали секции, занимавшиеся в том числе распознаванием образов и моделированием мышления, а среди известных отечественных исследователей ИИ энциклопедия называет также Г. С. Поспелова, Ю. И. Журавлёва и О. И. Ларичева [2].

К концу 1960-х годов стало ясно, что задачи, легко решаемые человеком, — понять фразу, узнать лицо, пройти по комнате — для машины гораздо сложнее шахмат. Программа ELIZA Джозефа Вейценбаума (1966) создавала иллюзию беседы с психотерапевтом, но не понимала ни слова. Её приём был прост: если собеседник писал, что ему грустно, программа могла спросить, почему ему грустно, то есть подставляла его же слова в заготовленный шаблон. В 1969 году Минский и Пейперт опубликовали книгу «Перцептроны», где показали ограничения однослойных нейронных сетей, и интерес к этому направлению надолго угас. Финансирование сократилось, и

в 1970-е годы наступила так называемая первая «зима» искусственного интеллекта.

Оглядываясь назад, можно сказать, что первые десятилетия ИИ были временем завышенных ожиданий. Однако именно тогда сформировались два больших подхода, спор которых определил дальнейшую историю. Символьный подход описывал знания с помощью правил и логики, коннекционистский пытался воспроизвести работу нейронных сетей мозга [5]. В разные периоды верх брал то один, то другой.

1.2 От экспертных систем к глубокому обучению

Выход из первой «зимы» обеспечил символьный подход. С середины 1970-х годов распространились экспертные системы — программы, которые хранили знания специалистов в виде правил «если — то» и делали на их основе выводы. Их применяли для генерации химических формул, медицинской диагностики, прогноза месторождений полезных ископаемых [2]. Широко известна система MYCIN, разработанная в Стэнфорде для диагностики бактериальных инфекций и подбора антибиотиков. В нашей стране создавались собственные экспертные системы, например для анализа текстов и оценки качества воды, а Д. А. Поспелов развивал методы ситуационного управления [2]. Правила такой системы сопровождались коэффициентами уверенности: программа не просто называла вероятного возбудителя, а сообщала, насколько надёжен её вывод.

В 1980-е годы экспертные системы стали коммерческим продуктом, а Япония в 1982 году объявила о проекте компьютеров пятого поколения, рассчитанных на логический вывод. Однако системы на правилах оказались хрупкими. Знания приходилось вручную извлекать у экспертов, база правил разрасталась до тысяч пунктов, и программа терялась в любой ситуации, не предусмотренной разработчиками. К концу десятилетия наступила вторая «зима» [4].

Параллельно возрождался коннекционизм. В 1986 году Д. Румельхарт, Дж. Хинтон и Р. Уильямс популяризировали метод обратного распространения ошибки, позволивший обучать многослойные нейронные сети. Этот алгоритм снимал ограничения, на которые указывали Минский и Пейперт. Тем не менее для настоящего прорыва не хватало двух вещей — больших объёмов данных и вычислительной мощности.

В 1990-е годы на первый план вышло машинное обучение в широком смысле. Его суть в том, что программа не получает готовых правил, а извлекает закономерности из примеров. Различают обучение с учителем, когда каждому примеру сопоставлен правильный ответ, обучение без учителя, когда система сама ищет структуру в данных, и обучение с подкреплением, при котором агент учится на наградах и штрафах за свои действия [4]. Метод опорных векторов, деревья решений, байесовские сети широко применялись для фильтрации спама, распознавания рукописных цифр, анализа банковских операций задолго до нынешней моды на нейросети.

Символическим событием 1990-х годов стала победа шахматного компьютера IBM Deep Blue над чемпионом мира Гарри Каспаровым в мае 1997 года. Победа была достигнута скорее за счёт перебора огромного числа вариантов, чем за счёт «понимания» игры, но она изменила общественное восприятие. Машина обыграла лучшего человека в игре, долго считавшейся вершиной интеллектуальной деятельности.

Подлинная революция началась в 2010-е годы. Интернет дал огромные массивы размеченных данных, а графические процессоры, созданные для компьютерных игр, оказались идеальными для обучения нейросетей. В 2012 году сверточная сеть AlexNet, разработанная в группе Хинтона, с большим отрывом выиграла конкурс ImageNet по распознаванию изображений. Я. Лекун, Й. Бенжио и Дж. Хинтон в обзорной статье 2015 года в журнале Nature описали глубокое обучение как подход, при котором многослойные модели сами выучивают представления данных на нескольких уровнях

абстракции, и подвели итог его успехов в распознавании речи, изображений и обработке текста [7].

Конкурс ImageNet (ILSVRC) хорошо показывает масштаб перемены. Участникам предлагалось отнести фотографию к одной из тысячи категорий, причём ответ засчитывался, если верная категория попадала в пять самых вероятных вариантов программы. В 2010 и 2011 годах победители, опиравшиеся на традиционные методы, ошибались в 28,2% и 25,8% случаев. Команда SuperVision, представившая AlexNet, в 2012 году снизила ошибку до 16,4%, через год победитель показал 11,7%, а в 2014 году сеть GoogLeNet довела её до 6,7% [10]. Авторы обзора результатов конкурса подсчитали, что за пять лет ошибка сократилась в 4,2 раза (рисунок 2).

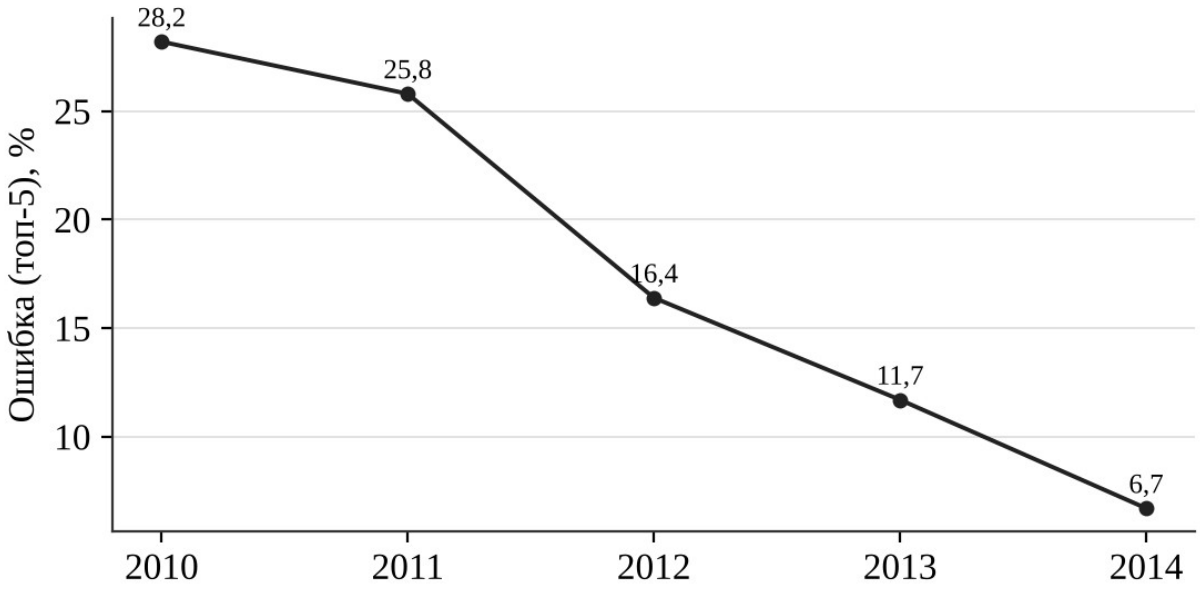


Рисунок 2 – Ошибка победителей конкурса ImageNet в классификации изображений, 2010–2014 гг.

Источник: [10].

Самый крутой участок кривой приходится на 2012 год, когда в соревнование вошли глубокие сверточные сети. С тех пор они стали стандартным инструментом компьютерного зрения.

Дальше события ускорились. В марте 2016 года программа AlphaGo компании DeepMind победила одного из сильнейших игроков в го Ли Седоля со счётом 4:1, хотя ещё за несколько лет до того такой результат считался

делом далёкого будущего. В 2017 году исследователи Google предложили архитектуру «трансформер» [13]. Её ключевой механизм — внимание: обрабатывая каждое слово, модель оценивает, какие другие слова фразы для него существенны, причём делает это для всех слов одновременно, а не по очереди. На трансформерах построены современные большие языковые модели. В ноябре 2022 года компания OpenAI открыла публичный доступ к ChatGPT, и генеративный ИИ за считанные месяцы вошёл в повседневную жизнь миллионов людей.

Признание пришло и от научного сообщества. В 2024 году Нобелевскую премию по физике получили Джон Хопфилд и Джеффри Хинтон за фундаментальные открытия, сделавшие возможным машинное обучение с помощью искусственных нейронных сетей, а премию по химии разделили Дэвид Бейкер, Демис Хассабис и Джон Джампер — двое последних за систему AlphaFold, предсказывающую структуру белков. Ещё недавно сложно было представить, что работы по ИИ будут отмечены высшими наградами в естественных науках.

История ИИ напоминает маятник: волны энтузиазма сменялись «зимами» разочарования, а ведущий подход не раз менялся (таблица 1).

Таблица 1 – Основные этапы развития искусственного интеллекта

Период	**Ведущий подход**	**Характерные события и системы**
1943–1956	Зарождение идей	Модель нейрона (1943), тест Тьюринга (1950), Дартмутский семинар (1956)
1956–1969	Символьный поиск, первые нейросети	«Логик-теоретик», Лисп, перцептрон, ELIZA
1970-е	Первая «зима»	Книга «Перцептроны», сокращение финансирования
Середина 1970-х — 1980-е	Экспертные системы	MYCIN, японский проект ЭВМ пятого поколения
Конец 1980-х — 2000-е	Статистическое машинное обучение	Метод опорных векторов, Deep Blue (1997)
2012–2016	Глубокое обучение	AlexNet (2012), AlphaGo (2016)
С 2017 г.	Трансформеры, генеративный ИИ	ChatGPT (2022), AlphaFold, Нобелевские премии (2024)

Источник: составлено автором по [2; 4; 7; 10].

Нынешний подъём отличается от прежних тем, что технологии уже приносят ощутимый экономический результат. Это не гарантирует отсутствия новых спадов, но делает их менее вероятными, чем в эпоху, когда ИИ существовал в основном в лабораториях.

2. Области применения искусственного интеллекта

Сегодня трудно назвать отрасль, где ИИ не нашёл бы применения. Рассмотрим наиболее показательные из них.

2.1 Искусственный интеллект в экономике, медицине и транспорте

Масштаб вложений в ИИ поражает. По данным AI Index 2026, совокупные корпоративные инвестиции в искусственный интеллект в 2025 году достигли 581,69 млрд долларов, увеличившись за год на 129,9% [6]. Распределение крайне неравномерно: частные инвестиции в США составили почти 285,9 млрд долларов, что в 23 раза больше, чем в следующем по объёму Китае, где они равнялись 12,4 млрд (рисунок 3).

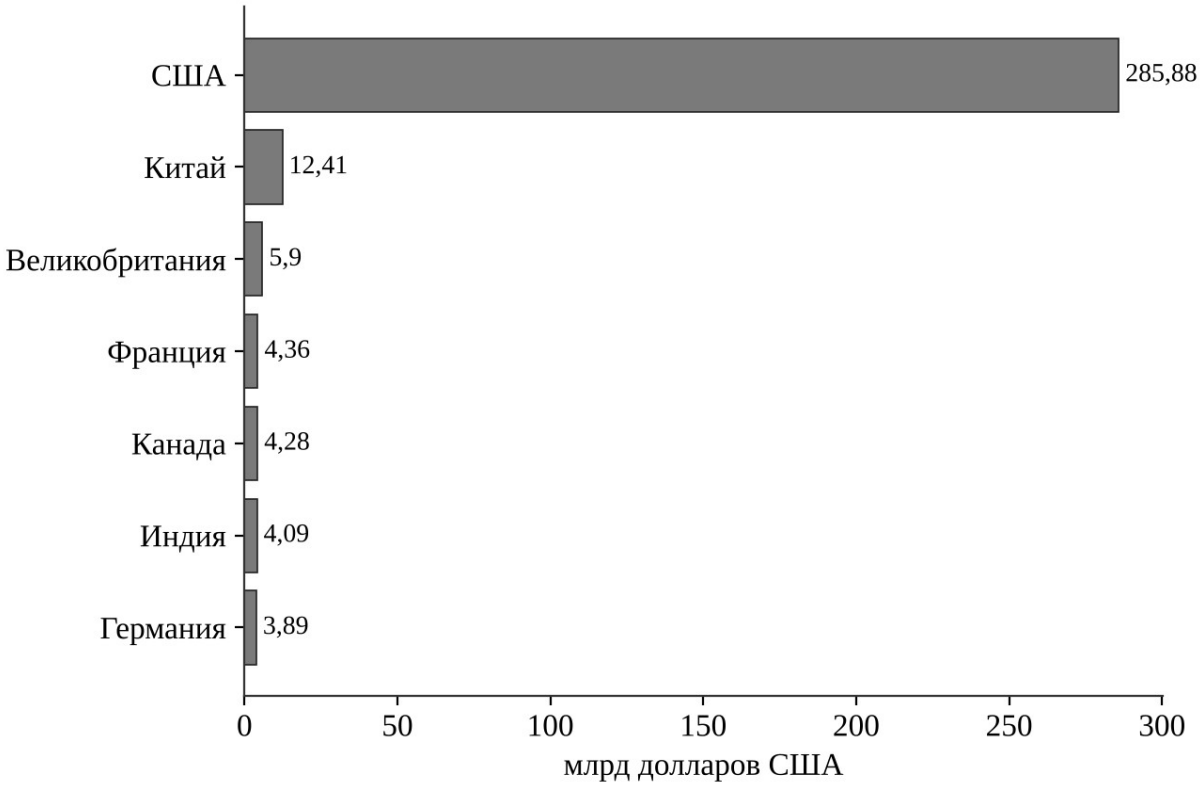


Рисунок 3 – Частные инвестиции в искусственный интеллект по странам в 2025 году

Источник: [6].

Следом за Китаем идут Великобритания, Франция и Канада с объёмами от 4 до 6 млрд долларов [6]. Шесть стран, показанных на рисунке после США, в сумме привлекли около 35 млрд долларов, то есть примерно

восьмую часть американского объёма. Авторы доклада оговариваются, что китайские расходы, вероятно, занижены из-за роли государственных фондов [6].

Бизнес использует ИИ прежде всего там, где много однотипных операций и данных. Банки оценивают кредитоспособность заёмщиков, выявляют мошеннические транзакции, отвечают клиентам через чат-ботов. Торговые сети прогнозируют спрос и управляют запасами, промышленные компании с помощью датчиков и моделей предсказывают поломки оборудования до того, как они случились. Доля организаций, регулярно применяющих генеративный ИИ хотя бы в одной функции, выросла за год с 71% до 79% [6].

Каков реальный эффект для производительности? Здесь картина неоднозначна. Исследования, обобщённые в AI Index, дают заметно различающиеся оценки (таблица 2).

Таблица 2 – Влияние ИИ-ассистентов на производительность труда

Категория работников	**Изменение показателя**
Операторы службы поддержки	+14–15% решённых обращений в час
Разработчики программ	+26% выполненных запросов на изменение кода
Маркетологи	+50% выработки
Опытные разработчики открытого ПО (эксперимент METR)	−19% скорости выполнения задач

Источник: составлено автором по [6].

Выигрыш сильнее всего заметен там, где ИИ-помощник подсказывает типовой ответ клиенту, дописывает шаблонный код или готовит черновик рекламного текста [6]. Эксперимент организации METR дал противоположный результат: опытные разработчики открытого программного обеспечения при использовании ИИ-ассистента работали на 19% медленнее. Иначе говоря, ИИ хорошо ускоряет типовую работу, но на сложных задачах, где человек и так прекрасно знает материал, может мешать: часть времени уходит на проверку и исправление подсказок.

В медицине ИИ особенно успешен в анализе изображений. Нейросети находят признаки опухолей на маммограммах и снимках лёгких, выявляют диабетическую ретинопатию по фотографиям глазного дна, помогают патоморфологам анализировать срезы тканей. Важно понимать, что речь идёт о системах поддержки решений: окончательный диагноз ставит врач, а программа служит вторым мнением, которое не устаёт к концу смены. В Москве с 2020 года идёт эксперимент по применению компьютерного зрения в лучевой диагностике: алгоритмы просматривают снимки, поступающие из городских поликлиник, и отмечают участки, на которые рентгенологу стоит обратить внимание.

Отдельная история — фундаментальная биология. Чтобы экспериментально определить структуру одного белка, требуются месяцы, а то и годы кропотливой работы, поэтому к началу 2020-х годов была известна структура примерно 100 тысяч уникальных белков при миллиардах известных белковых последовательностей [12]. Система AlphaFold научилась предсказывать трёхмерную структуру белка по последовательности аминокислот. На конкурсе CASP14 в 2020 году её новая версия показала точность, в большинстве случаев сопоставимую с экспериментальными методами [12]. Задача, над которой учёные бились полвека, была в значительной мере решена, а базы предсказанных структур стали бесплатно доступны исследователям по всему миру. Это ускоряет разработку лекарств и изучение болезней.

Транспорт — третья большая сфера. Системы помощи водителю, удерживающие автомобиль в полосе и тормозящие перед препятствием, стали обычной опцией. В декабре 2018 года компания Waymo запустила в американском Финиксе первый коммерческий сервис беспилотного такси, поначалу с инженером-наблюдателем на водительском месте. Сейчас роботакси возят пассажиров в нескольких городах США и Китая, а в России испытания беспилотных такси проводятся в Москве. Логистические компании оптимизируют маршруты доставки, авиакомпании — расписания и

загрузку самолётов. Полностью автономное вождение в любых условиях пока остаётся нерешённой задачей, но частичная автоматизация уже экономит время и топливо.

Ещё одна область, где ИИ изменил повседневность, — работа с языком. Машинный перевод прошёл путь от программ на правилах, выдававших курьёзные фразы, к нейросетевым системам, которые переводят деловую переписку и техническую документацию на вполне приемлемом уровне. Распознавание речи позволяет диктовать сообщения, получать расшифровку лекций и звонков, управлять устройствами голосом. Для людей с нарушениями зрения или слуха такие технологии становятся средством доступа к информации.

В России крупнейшими разработчиками ИИ выступают технологические компании и банки. Отечественные большие языковые модели, такие как YandexGPT и GigaChat, появились в 2023 году, а голосовые помощники стали привычной частью смартфонов и умных колонок. Применение ИИ на производстве, в сельском хозяйстве и энергетике растёт медленнее, чем в финансах и торговле, что связано с меньшим объёмом накопленных цифровых данных и нехваткой специалистов.

Подводя итог, экономический эффект ИИ уже заметен, но распределён неравномерно: он выше там, где задачи структурированы, а данных много. Именно поэтому цифровизация процессов предшествует внедрению ИИ, а не следует за ним.

2.2 Образование, наука и государственное управление

В образовании ИИ используется для персонализации обучения, автоматической проверки заданий и учебной аналитики. Е. Другова и её соавторы проанализировали 43 зарубежные статьи о применении ИИ для учебной аналитики в высшей школе и соотнесли решения с этапами модели педагогического проектирования ADDIE. Оказалось, что почти все

разработки относятся к этапу реализации курса, а для этапов анализа и проектирования нашлись лишь одна и две системы соответственно [3]. Авторы объясняют перекос тем, что данные о студентах появляются только после начала курса.

Для студентов главной новостью последних лет стали чат-боты. Они помогают разобраться в сложной теме, найти ошибку в коде, перевести статью. Но та же технология позволяет сгенерировать готовый реферат, и вузы вынуждены пересматривать формы контроля, смещая акцент на устные защиты, проекты и работу в аудитории. Авторы упомянутого обзора подчёркивают, что прогресс невозможен без развития компетенций самих преподавателей в области ИИ и учебной аналитики [3].

Появились и сервисы, которые пытаются использовать чат-боты в учебных целях, а не в обход учёбы. Пример — Khanmigo, помощник некоммерческой образовательной платформы Khan Academy, представленный в 2023 году на основе модели GPT-4. Он задуман как репетитор: вместо готового ответа задаёт наводящие вопросы и просит ученика объяснить ход решения.

В науке ИИ превращается в инструмент открытий. Помимо биологии, нейросети используют для поиска новых материалов, обработки данных телескопов и ускорителей частиц, прогноза погоды. Модели, обученные на архивах наблюдений, уже конкурируют с традиционными численными прогнозами по точности на несколько дней вперёд, требуя при этом значительно меньше вычислений.

Государственное управление использует ИИ для обработки обращений граждан, анализа налоговой отчётности, управления дорожным движением и видеонаблюдения. В Москве в 2020 году специальным федеральным законом был введён пятилетний экспериментальный правовой режим, упрощающий разработку и внедрение технологий ИИ, в том числе работу с обезличенными персональными данными. Ошибка алгоритма, распределяющего пособия или выявляющего нарушителей, затрагивает права граждан, поэтому вместе с

внедрением технологий государства вырабатывают правила их использования, о которых пойдёт речь в третьей главе.

3. Принципы работы и границы возможностей искусственного интеллекта

Рассказ о достижениях ИИ остаётся неполным без понимания того, как такие системы получают свои умения и почему иногда ошибаются. В этой главе на простом числовом примере разобрано обучение нейронной сети, а затем рассмотрены ограничения современных систем, связанные с ними риски и ответ государств на эти риски.

3.1 Как обучается нейронная сеть: разбор простого примера

Искусственный нейрон устроен проще, чем можно подумать. Он получает на вход несколько чисел, умножает каждое на свой вес, складывает результаты и пропускает сумму через несложную функцию, определяющую силу выходного сигнала. Веса и есть «знания» сети. Перцептрон Розенблатта состоял из таких элементов, а в современных глубоких сетях их соединяют в десятки и сотни слоёв, так что выход одного слоя служит входом для следующего [5].

Обучить сеть — значит подобрать веса так, чтобы её ответы на обучающих примерах как можно меньше отличались от правильных. Расхождение измеряют функцией ошибки. Затем вычисляют, в какую сторону нужно изменить каждый вес, чтобы ошибка уменьшилась, и сдвигают его на небольшой шаг. Этот приём называют градиентным спуском [7]. Если представить ошибку как холмистую местность, алгоритм всё время делает шаг вниз по склону, пока не окажется в низине.

Разберём процедуру на самом простом случае — модели с одним весом. Пусть известно, что пешеход за 2 часа прошёл 6 км, а модель должна найти его скорость w. Прогноз модели равен 2w, ошибка — разности между прогнозом и правильным ответом, а функцией ошибки служит квадрат этой разности. Производная квадрата ошибки по весу равна $4 \times (2w - 6)$. Правило обновления такое: из текущего веса вычитается эта производная, умноженная

на скорость обучения. Возьмём начальный вес 0 и скорость обучения 0,05. Результаты первых шагов приведены в таблице 3.

Таблица 3 – Подбор веса модели методом градиентного спуска

Шаг	Вес w, км/ч	Прогноз 2w, км	Ошибка, км
0	0	0	−6,00
1	1,20	2,40	−3,60
2	1,92	3,84	−2,16
3	2,35	4,70	−1,30
4	2,61	5,22	−0,78
5	2,77	5,53	−0,47
10	2,98	5,96	−0,04

Источник: расчёт автора.

На первом шаге производная равна 4 × (0 − 6) = −24, и вес увеличивается на 0,05 × 24 = 1,2. Дальше ошибка на каждом шаге сокращается на одну и ту же долю — 40%, и уже к десятому шагу модель почти нашла правильный ответ: 3 км/ч. Выбор скорости обучения тоже имеет значение. При скорости 0,3 вес начал бы перескакивать через правильное значение, а ошибка — расти от шага к шагу. Слишком маленькая скорость, наоборот, сделала бы обучение мучительно долгим.

В настоящей нейросети происходит то же самое, только весов не один, а миллионы и миллиарды. Уже в 2015 году Я. Лекун, Й. Бенжио и Дж. Хинтон писали, что в типичной системе глубокого обучения могут быть сотни миллионов настраиваемых весов и сотни миллионов размеченных примеров для обучения [7]. Вычислить, как ошибка зависит от каждого веса, позволяет метод обратного распространения ошибки, упомянутый в первой главе. Сигнал ошибки проходит по сети от выхода к входу, и каждый слой получает свою «долю вины». Веса обычно обновляют не по всему массиву данных сразу, а по небольшим случайным порциям примеров. Такой вариант называют стохастическим градиентным спуском [7].

Подбор весов составляет лишь один этап более длинного процесса создания системы машинного обучения (рисунок 4).



Рисунок 4 – Этапы создания системы машинного обучения

Источник: составлено автором по [7].

Проверка на отдельной тестовой выборке нужна потому, что модель способна просто запомнить обучающие примеры. О качестве системы судят по тому, как она справляется с примерами, которых не видела при обучении [7]. Если на знакомых данных ошибка почти нулевая, а на новых высокая, говорят о переобучении. Модель в этом случае похожа на студента, который выучил ответы к билетам, но не понял предмет.

Большие языковые модели обучаются по тому же принципу, хотя задача у них особая: предсказать следующий фрагмент текста по предыдущим. Правильный ответ содержится в самом тексте, поэтому ручная разметка не нужна, и модель можно учить на огромных массивах книг, статей и веб-страниц. Архитектура трансформер позволила эффективно распараллелить такие вычисления [13].

Из разобранного примера вытекает существенный вывод. Нейросеть не понимает мир в человеческом смысле: она находит веса, при которых ошибка на обучающих данных мала. Если в данных есть пробелы или перекосы, модель их унаследует, а на примерах, непохожих на обучающие, может ошибиться совершенно непредсказуемо.

3.2 Ограничения, риски и правовое регулирование

Риски вполне реальны. По данным AI Index 2026, число зарегистрированных инцидентов, связанных с ИИ, в 2025 году достигло рекордных 362 [6]. Среди проблем — дискриминация, когда модель, обученная на исторических данных, воспроизводит заложенные в них предубеждения; дипфейки, используемые для мошенничества; «галлюцинации» языковых моделей, уверенно сообщающих несуществующие факты.

Каждая из этих угроз уже проявилась на практике. В 2018 году стало известно, что компания Amazon отказалась от экспериментальной системы отбора резюме. Её обучали на заявках прошлых лет, когда среди кандидатов на технические должности преобладали мужчины, и система стала занижать оценку резюме, где встречалось слово «женский», например в названии студенческого клуба. В 2023 году федеральный суд в Нью-Йорке оштрафовал двух адвокатов на 5 тыс. долларов: в поданном ими документе были ссылки на несуществующие судебные решения, придуманные ChatGPT. Для студентов вывод очевиден: ссылку или цитату, предложенную чат-ботом, необходимо проверять по первоисточнику.

Серьёзная проблема — непрозрачность. Глубокая нейросеть содержит миллионы или миллиарды параметров, и даже её создатели не всегда могут объяснить, почему модель приняла то или иное решение. Для рекомендаций фильма это неважно, для отказа в кредите или медицинского заключения — принципиально. Отсюда интерес к объяснимому ИИ и требования регуляторов раскрывать логику работы систем в чувствительных сферах. Обучение крупных моделей требует и огромных вычислительных ресурсов, а значит, электроэнергии, что придаёт развитию ИИ экологическое измерение.

Есть и ограничения, которые не сводятся к качеству данных. Модель уверенно работает в пределах того, что видела при обучении, но плохо переносит знания на непохожие ситуации. Незаметное человеку изменение нескольких пикселей способно заставить сеть распознавания изображений ошибиться, а языковая модель может справиться со сложной задачей и

запутаться в простой, если та сформулирована непривычно. Отсюда осторожность при использовании ИИ там, где ошибка стоит дорого: в медицине, на транспорте, в суде.

Наконец, остаётся вопрос о рынке труда. ИИ автоматизирует не столько профессии целиком, сколько отдельные задачи внутри них. Переводчик, юрист или программист не исчезают, но их работа меняется: рутинную часть берёт на себя машина, а от человека требуются проверка, постановка задачи и ответственность за результат. Философы и инженеры давно различают «сильный» ИИ, равный человеческому разуму во всей его широте, и «слабый», решающий конкретные задачи. Все современные системы, при всей их впечатляющей мощи, относятся ко второму типу.

Государства стремятся одновременно поддержать развитие ИИ и ограничить риски. В России в 2019 году указом Президента № 490 была утверждена Национальная стратегия развития искусственного интеллекта на период до 2030 года. Стратегия определяет ИИ как 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 [1]. В феврале 2024 года указ № 124 обновил документ: появилось понятие доверенных технологий ИИ, а также задача их обязательного внедрения в областях, где возможен ущерб безопасности страны [1]. Действует и национальный стандарт ГОСТ Р 59277–2020, задающий классификацию систем ИИ [2].

Обновлённая стратегия содержит измеримые ориентиры на 2030 год, часть которых приведена в таблице 4.

Таблица 4 – Некоторые целевые показатели Национальной стратегии развития искусственного интеллекта

Показатель	**2022 г.**	**2030 г. (цель)**
Объём услуг в области ИИ, млрд руб. в год	12	не менее 60
Выпускники вузов по направлениям ИИ, тыс. чел. в год	около 3	15,5
Доверие граждан к технологиям ИИ, %	55	не менее 80
Приоритетные отрасли с высокой готовностью к внедрению ИИ, %	12	95

Источник: составлено автором по [1].

Самый амбициозный ориентир касается готовности отраслей: их долю предстоит увеличить почти в восемь раз. Объём услуг и выпуск специалистов должны вырасти примерно впятеро.

Европейский союз пошёл по пути жёсткого регулирования. Регламент об искусственном интеллекте, вступивший в силу 1 августа 2024 года, делит системы по уровням риска и для каждого уровня устанавливает свои требования (рисунок 5) [11].

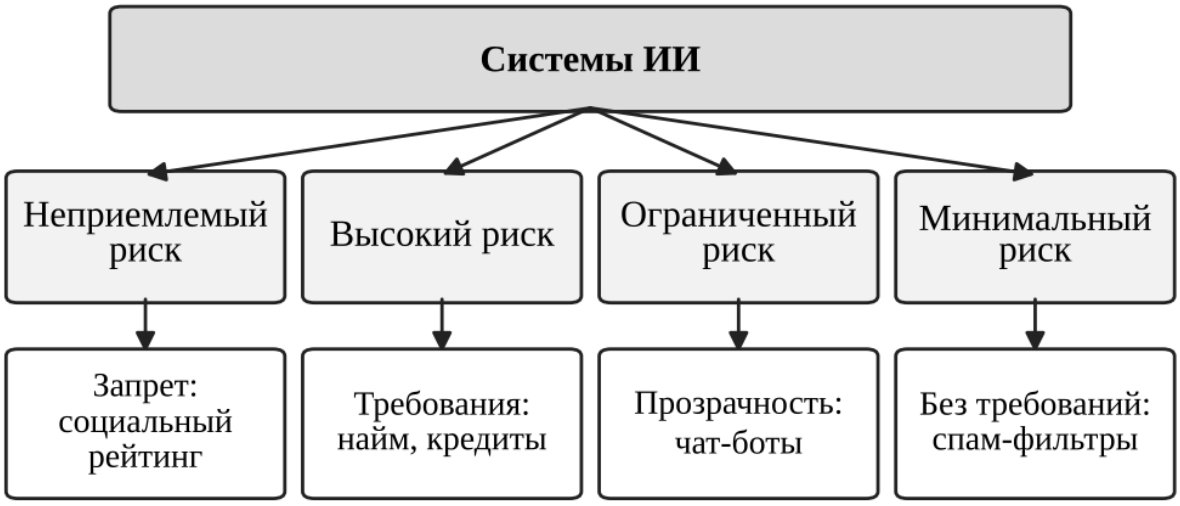


Рисунок 5 – Уровни риска систем искусственного интеллекта по Регламенту ЕС

Источник: составлено автором по [11].

Запреты на практики неприемлемого риска, например на социальный рейтинг граждан, применяются с февраля 2025 года. К системам высокого риска относятся, в частности, программы отбора резюме, оценки кредитоспособности и проверки экзаменационных работ: их разработчики обязаны управлять рисками, документировать работу системы и обеспечивать надзор человека. Для чат-ботов и генераторов изображений главное требование — прозрачность: пользователь должен знать, что общается с машиной, а дипфейки должны быть помечены [11]. Большинство же систем, вроде спам-фильтров и компьютерных игр, особым требованиям не подлежат.

В России в 2021 году крупные компании подписали Кодекс этики в сфере ИИ, то есть сделали ставку на саморегулирование. Какой путь эффективнее, покажет время.

Сам тест Тьюринга, кстати, сегодня выглядит не так однозначно, как в 1950 году. Языковые модели способны вести беседу, неотличимую от человеческой, но означает ли это мышление? Спор, начатый Тьюрингом, продолжается, только теперь он ведётся не о гипотетических машинах, а о программах, которыми пользуются сотни миллионов людей.

ЗАКЛЮЧЕНИЕ

Проведённый анализ позволяет ответить на поставленные вопросы. Искусственный интеллект как научное направление оформился в середине XX века: в 1950 году А. Тьюринг сформулировал вопрос о мыслящих машинах, в 1955–1956 годах авторы Дартмутского проекта дали направлению имя, а первые программы доказывали теоремы и играли в шахматы. Советские учёные внесли заметный вклад в теорию машинного обучения и компьютерные шахматы.

Развитие ИИ шло волнами. Экспертные системы 1970–1980-х годов уступили место машинному обучению, а в 2010-е годы благодаря большим данным и графическим процессорам произошёл прорыв глубокого обучения. Его наглядно показывает конкурс ImageNet: за 2010–2014 годы ошибка лучших программ распознавания изображений снизилась с 28,2 до 6,7%. Затем появились трансформеры и большие языковые модели, а работы по ИИ были отмечены Нобелевскими премиями 2024 года.

Сегодня ИИ применяется в финансах, торговле, промышленности, медицине, транспорте, образовании, науке и государственном управлении; в 2025 году его использовали 88% организаций, охваченных исследованием AI Index. Эффект для производительности заметен в типовых задачах, но на сложных работах может оказаться и отрицательным.

Разбор простого примера показал, что обучение нейросети сводится к постепенному подбору весов, уменьшающему ошибку на примерах. Отсюда вытекают и сильные стороны ИИ, и его слабости: модель хороша там, где данных много и они представительны, но наследует их перекосы, может «галлюцинировать» и с трудом объясняет свои решения. Россия закрепила в Национальной стратегии конкретные ориентиры до 2030 года и ввела понятие доверенных технологий ИИ, а Европейский союз выбрал регулирование по уровням риска.

Разумное отношение к ИИ состоит в том, чтобы видеть в нём мощный инструмент, требующий от человека знаний и критического мышления. Для

будущего специалиста это означает необходимость понимать хотя бы основы того, как устроены и где ошибаются системы, которыми он пользуется каждый день.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. О внесении изменений в Указ Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» и в Национальную стратегию, утвержденную этим Указом : Указ Президента Российской Федерации от 15.02.2024 № 124 // Собрание законодательства Российской Федерации. – 2024. – № 8. – Ст. 1102.

2. Искусственный интеллект. – Текст : электронный // Большая российская энциклопедия : научно-образовательный портал. – URL: https://bigenc.ru/c/iskusstvennyi-intellekt-ac9fb0 (дата обращения: 22.10.2026).

3. Искусственный интеллект для учебной аналитики и этапы педагогического проектирования: обзор решений / Е. Другова, И. Журавлева, У. Захарова [и др.] // Вопросы образования. – 2022. – № 4. – С. 107–153. – DOI 10.17323/1814-9545-2022-4-107-153.

4. Рассел, С. Искусственный интеллект. Современный подход. Т. 1. Решение проблем. Знания и рассуждения / С. Рассел, П. Норвиг ; пер. с англ. А. В. Слепцова. – 4-е изд. – Москва : Вильямс, 2021. – 704 с. – ISBN 978-5-907365-25-4.

5. Ясницкий, Л. Н. Введение в искусственный интеллект : учебное пособие / Л. Н. Ясницкий. – 3-е изд., стер. – Москва : Академия, 2010. – 176 с. – ISBN 978-5-7695-7042-1.

6. Artificial Intelligence Index Report 2026 / Stanford Institute for Human-Centered Artificial Intelligence. – Stanford : Stanford University, 2026. – Текст : электронный. – URL: https://hai.stanford.edu/ai-index/2026-ai-index-report (дата обращения: 22.10.2026).

7. LeCun, Y. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // Nature. – 2015. – Vol. 521, № 7553. – P. 436–444. – DOI 10.1038/nature14539.

8. McCarthy, J. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955 / J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon // AI Magazine. – 2006. – Vol. 27, № 4. – P. 12–14.

9. Turing, A. M. Computing Machinery and Intelligence / A. M. Turing // Mind. – 1950. – Vol. 59, № 236. – P. 433–460.

10. Russakovsky, O. ImageNet Large Scale Visual Recognition Challenge / O. Russakovsky, J. Deng, H. Su [et al.] // International Journal of Computer Vision. – 2015. – Vol. 115, № 3. – P. 211–252. – DOI 10.1007/s11263-015-0816-y.

11. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). – Текст : электронный // Official Journal of the European Union. – 2024. – L series, 12.07.2024. – URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj (дата обращения: 22.10.2026).

12. Jumper, J. Highly accurate protein structure prediction with AlphaFold / J. Jumper, R. Evans, A. Pritzel [et al.] // Nature. – 2021. – Vol. 596, № 7873. – P. 583–589. – DOI 10.1038/s41586-021-03819-2.

13. Vaswani, A. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar [et al.] // Advances in Neural Information Processing Systems 30 (NIPS 2017). – 2017. – P. 5998–6008.