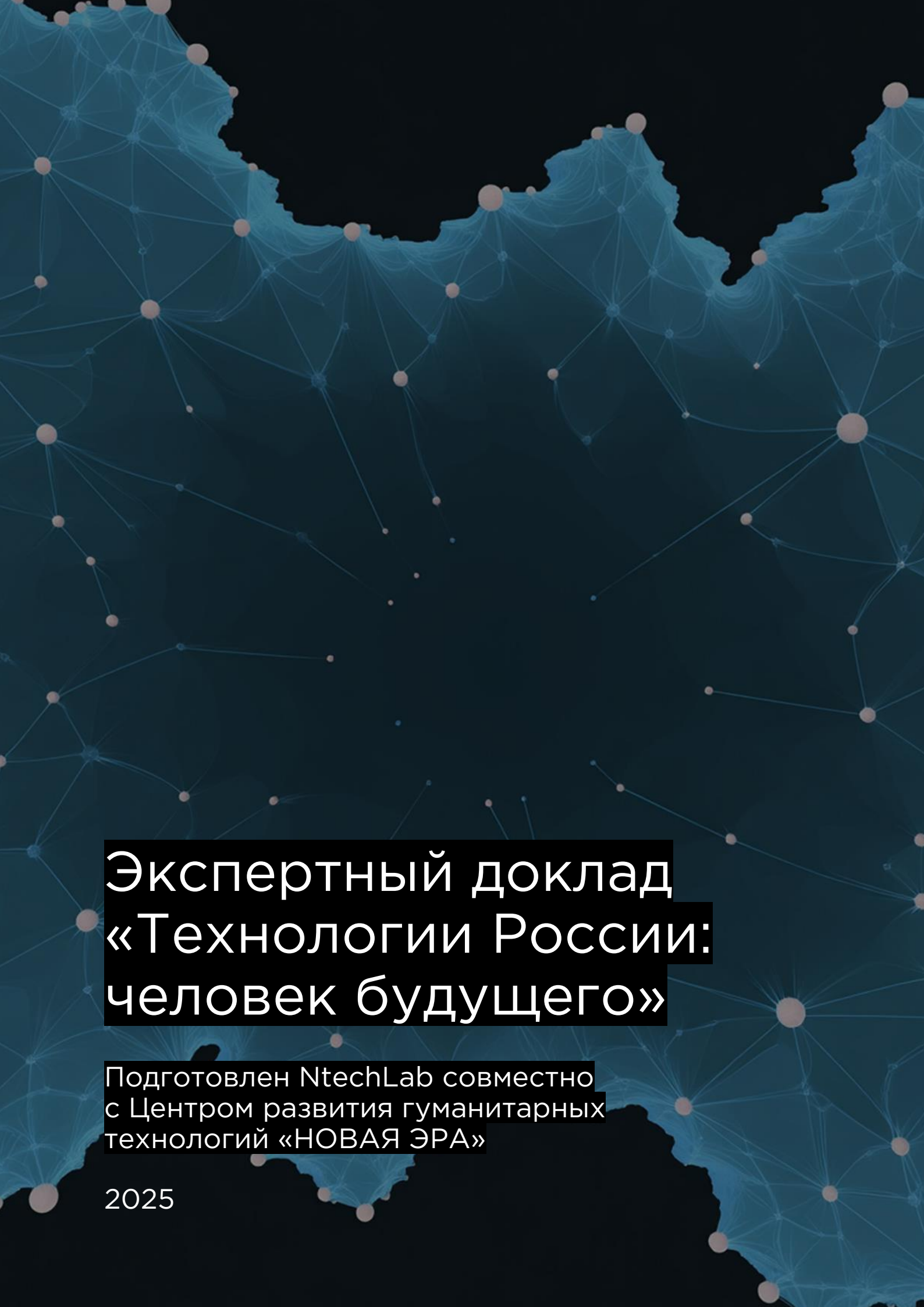




Экспертный доклад «Технологии России: человек будущего»

Подготовлен NtechLab совместно
с Центром развития гуманитарных
технологий «НОВАЯ ЭРА»

2025

Авторы доклада

Антюшин Максим Владимирович

Начальник управления цифровых технологий АНО «Горький Тех»

Бровко Василий Юрьевич

Директор по особым поручениям Госкорпорации Ростех

Войнов Артём Александрович

Менеджер интеллектуальных систем АНО «Горький Тех»

Володенков Сергей Владимирович

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Гнидко Константин Олегович

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Кайдин Ян Дмитриевич

Эксперт Центра развития «НОВАЯ ЭРА»

Карапетян Грант Каренович

Руководитель направления по работе с партнерами и развитию партнерской сети АНО «Горький Тех»

Красильников Илья Олегович

Проректор по цифровому развитию Московского политехнического университета

Кузнецов Андрей Владимирович

Директор лаборатории FusionBrain Института AIRI, кандидат технических наук

Любченко Александр Александрович

Заведующий лабораторией искусственного интеллекта, ООО «ОЦРВ», филиал «Сириус», холдинг «РЖД-Технологии»

Магси Назир Заман

Преподаватель кафедры менеджмента Московского политехнического университета

Авторы доклада

Митюков Андрей Алексеевич

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Назаров Александр Юрьевич

Заместитель генерального директора Госкорпорации Ростех

Никанов Иван Андреевич

Заместитель руководителя Исследовательского центра в сфере ИИ Университета Иннополис

Паламарчук Алексей Степанович

Генеральный директор NtechLab

Пивень Ольга Игоревна

Сооснователь и директор конференции ЦИПР

Попов Павел Александрович

Заместитель генерального директора «НИИАС РЖД»

Сальманов Олег Робертович

Основатель и редактор Telegram-канала «Нецифровая экономика»

Оглавление

Блок 1. Искусственный интеллект и человек	6
<u>РАЗГРАНИЧЕНИЕ ОБЯЗАННОСТЕЙ ЧЕЛОВЕКА И ИИ</u>	<u>7</u>
<u>ЭТИЧЕСКИЕ ГРАНИЦЫ ИСПОЛЬЗОВАНИЯ ИИ В ПОВСЕДНЕВНОЙ ЖИЗНИ</u>	<u>14</u>
<u>ЭТИКА И СДЕРЖИВАНИЯ В РАЗВИТИИ ИИ.....</u>	<u>20</u>
<u>Влияние ИИ на личность, мышление и поведение человека.....</u>	<u>26</u>
<u>ПОТРЕБНОСТЬ В ЭТИЧЕСКИХ КОДЕКСАХ ИЛИ ПРАВОВЫХ НОРМАХ</u>	<u>33</u>
<u>ИИ КАК СОБЕСЕДНИК И ДРУГ ДЛЯ МОЛОДЕЖИ</u>	<u>38</u>
<u>ИИ КАК ПОМОЩНИК В УЧЕБЕ И РАБОТЕ ДЛЯ МОЛОДЕЖИ.....</u>	<u>43</u>
<u>РИСК БЕЗРАБОТИЦЫ ИЗ-ЗА ИИ</u>	<u>48</u>
<u>ИИ — НОВЫЙ ПАРОВОЙ ДВИГАТЕЛЬ</u>	<u>54</u>
 Блок 2. Искусственный интеллект и страна	61
<u>НАИБОЛЕЕ ВОСТРЕБОВАННЫЕ ФУНКЦИИ ИИ ДЛЯ ГОС. ВЛАСТИ И БИЗНЕСА</u>	<u>62</u>
<u>ОСНОВНЫЕ ТРУДНОСТИ РЕГИОНОВ РФ ПРИ ВНЕДРЕНИИ ИИ</u>	<u>65</u>
<u>НАИБОЛЕЕ УСПЕШНЫЕ РЕГИОНЫ РОССИИ С ТОЧКИ ЗРЕНИЯ РАЗВИТИЯ ИИ</u>	<u>69</u>
<u>РОССИЙСКОЕ ЗАКОНОДАТЕЛЬСТВО И ТЕМПЫ РАЗВИТИЯ ИИ.....</u>	<u>74</u>
<u>ИЗМЕНЕНИЯ В СИСТЕМЕ ВЗАИМОДЕЙСТВИЯ РАЗНЫХ АКТОРОВ В ОБЛАСТИ ИИ.....</u>	<u>82</u>
<u>Ключевые достижения ИИ в России.....</u>	<u>86</u>
 Блок 3. Искусственный интеллект и будущее	91
<u>Ключевые функции ИИ ЧЕРЕЗ 10–15 ЛЕТ</u>	<u>92</u>
<u>РИСКИ ДЛЯ ЧЕЛОВЕЧЕСТВА ОТ ШИРОКОГО РАСПРОСТРАНЕНИЯ ИИ</u>	<u>109</u>
<u>ФАКТОРЫ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОЙ ИНТЕГРАЦИИ ИИ В ЖИЗНЬ ОБЩЕСТВА</u>	<u>116</u>
<u>НАДЕЛЕНИЕ ИИ ЭМПАТИЕЙ И ДУШОЙ.....</u>	<u>120</u>
<u>ПРОТИВОСТОЯНИЕ МЕЖДУ ЧЕЛОВЕКОМ И ИИ</u>	<u>126</u>
<u>ИИ КАК ФАКТОР КОНТРОЛЯ ЗА БЕЗОПАСНОСТЬЮ В ГОРОДАХ</u>	<u>131</u>

Введение

NtechLab в партнерстве с Центром развития гуманитарных технологий «НОВАЯ ЭРА» к завершению юбилейной конференции ЦИПР провели социологическое исследование среди молодежи России, чтобы выяснить, как современное молодое поколение воспринимает технологии искусственного интеллекта и как видит их дальнейшее развитие и применение в жизни общества. Данный экспертный доклад включает интерпретацию ряда результатов исследования и восприятие искусственного интеллекта людьми, которые вносят свой вклад в развитие технологий в России.

2025 год.

Блок 1

Искусственный интеллект и человек



**ВОПРОС**

Какие задачи, на ваш взгляд, допустимо полностью передать ИИ, а какие должны оставаться исключительно в зоне ответственности человека?

Ключевые выводы:**1. Непрозрачность алгоритмов ИИ.**

Описание: алгоритмы ИИ часто работают как «черный ящик», что делает их логику и принципы принятия решений непонятными для пользователей и общества. Необходимо обеспечить прозрачность алгоритмов, их логики и ценностей перед внедрением. Проводить общественное обсуждение и одобрение.

2. Противоречие между эффективностью и справедливостью.

Описание: решения ИИ могут быть оптимальными с точки зрения эффективности, но несправедливыми с человеческой точки зрения. Вопрос: какие принципы справедливости заложены в алгоритм? Необходимо четко определять ценности и принципы, на которых основан алгоритм (например, по Аристотелю). Проверять соответствие решений ИИ человеческим нормам. Необходима этическая настройка ИИ, чтобы его решения соответствовали общепринятым ценностям и видам.

3. Отсутствие контроля над алгоритмами со стороны общества.

Описание: решения часто принимаются «сверху» без информирования тех, кого они касаются. Люди не понимают, на каких принципах работает ИИ, что ведет к делегитимации технологий. Необходимо обеспечивать открытость процессов разработки и внедрения ИИ.

4. Процедуры внедрения ИИ.

Описание: процесс передачи управления ИИ должен быть прозрачным, поэтапным и одобренным обществом. Недостаточная проработанность этих процедур может привести к недоверию и рискам. Необходимо

разрабатывать прозрачные и согласованные с обществом процедуры внедрения ИИ. Устранять все сомнения до запуска. Сохранять возможность для остановки деятельности ИИ.

5. Ограниченная надежность ИИ в комплексных задачах и вопросах, связанных с потребностью в суждении и ответственности.

Описание: ИИ, пытаясь решать широкий круг задач одновременно, демонстрирует низкую надежность в сравнении с узкоспециализированными программами. Применение ИИ оправдано и необходимо там, где цена ошибки не является критичной для человека и общества. Необходимо разделять задачи: ИИ — для вспомогательных функций, человек — для контроля и финальных решений. Применение концепции Human-in-The-Loop.



Сергей Владимирович Володенков

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Не вижу задачи, которые необходимо полностью отдавать искусственному интеллекту, точно так же, как не вижу задачи, которые человек должен решать принципиально самостоятельно, без задействования каких-то вспомогательных систем.

Все зависит от того, как настроен алгоритм. В любом случае, когда мы что-то отдаем на откуп искусственному интеллекту, за ним все равно стоит разработчик алгоритма, на основе которого искусственный интеллект и работает. Ведь неслучайно Аниш Аниш разработал концепцию алгократии, согласно которой нас ожидает в будущем власть алгоритмов. Алгоритмы пока что создают люди, искусственный интеллект их самостоятельно не продуцирует.

Понятно, что мы говорим сейчас о текущем уровне развития искусственного интеллекта, то есть о слабом ИИ. Если мы говорим об Artificial General Intelligence, общем интеллекте, или тем более об Artificial Super Intelligence — сильном интеллекте,

который наверняка будет способен создавать самостоятельно алгоритмы, то тогда, с учетом того, что алгоритм — это «черный ящик», возникает вопрос, а на основе каких параметров, принципов и ценностей этот алгоритм выполняет свою работу.

Таким образом, мы можем прийти к решениям, которые с позиции искусственного интеллекта будут оптимальными, но при этом для человека все будет иначе. Иными словами, прийти к противопоставлению эффективности (ИИ) и справедливости (человек).

Справедливый в каком измерении? То есть, по Аристотелю, какой из четырех видов справедливости он реализует? Эти ответы должны быть понятны, прозрачны и приняты, прежде чем алгоритму можно будет дать возможность что-то делать.

Поэтому, по большому счету, можно прийти к выводу, что можно что-то отдать на откуп, но перед этим алгоритм должен действительно получить статус прозрачного и понятного по своим параметрам и логике работы, по тем ценностям и смыслам, которые в него будут заложены, и только после одобрения теми, по отношению к кому он будет применен.

Как правило, решения спускаются сверху, и, когда объект алгоритмического управления оказывается внутри такого процесса управления, как правило, он вообще не в курсе, на каких принципах им управляют. Есть декларируемые, а есть реальные принципы алгоритмического управления. Еще раз повторюсь, алгоритм «черного ящика» непрозрачный. Сегодня существует тренд на теневизацию алгоритмов искусственного интеллекта, и это может приводить к делегитимации процессов управления.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Любые вопросы, в которых ИИ доказанно оказывается лучше человека, стоит передавать ИИ. Другое дело, что у человека должна оставаться возможность «вырубить рубильник». Важно не то, где внедрять, а как именно: процедуры передачи управления в «руки» ИИ должны быть проверенными, поэтапными, прозрачными, явным и непротиворечивым образом одобрены обществом, сняты все консервы и т. д.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

На мой взгляд, все рутинные повторяющиеся задачи можно и нужно отдавать системам ИИ. В то же время в задачах, связанных с безопасностью, здоровьем на текущий момент требуются специальные процедуры доказательства безопасности работы. В настоящий момент общепринятым подходом для данной области является работа ИИ в роли экспертных систем, где окончательное решение принимает человек.



Андрей Владимирович Кузнецов

Директор лаборатории FusionBrain Института AIRI, кандидат технических наук

Безусловно, когда сегодня мы говорим об ассистентах, то в первую очередь мы должны научиться делегировать ИИ рутинные задачи, которые занимают у нас время, но выполняются регулярно, и никто за нас их сделать не может: ответить на письмо, извлечь из сообщения информацию и запланировать встречу в соответствии с имеющимся расписанием, дать первичный анализ повреждений по фото автомобиля для страховой и т.д. Какие-то задачи можно автоматизировать полностью (задачи первой линии поддержки), какие-то останутся на уровне верификации человеком (постановка точного диагноза по анамнезу). Таким образом, каждый обладатель такого помощника становится своего рода администратором: ставит задачу, получает ответ, дает ему оценку и верифицирует выполнение.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Технологии ИИ применяются уже давно, и множество задач давно делегированы специализированным программам. Однако, если говорить именно об искусственном интеллекте в современном его понимании, в наиболее популярных решениях, таких как OpenAI, DeepSeek и их аналоги, то полностью передать им какую-либо задачу на 100% невозможно.

Причина в том, что эти системы не обладают строго детерминированной логикой решения задач. Их подход к получению результата носит вероятностный и в определенной степени хаотичный характер, пусть даже с долей предсказуемости. Пытаясь решать широкий круг задач одновременно, они зачастую не способны гарантировать безошибочное выполнение даже одной конкретной задачи.

Можно возразить, ссылаясь на человеческий фактор. Однако ключевое отличие в том, что человек несет ответственность за свои действия и решения. Он способен объяснить ход своих мыслей, логику и мотивы, что делает его действия осмысленными и прозрачными. Именно поэтому критически важные задачи, особенно те, что связаны с моральным выбором, ответственностью и последствиями, должны оставаться в зоне ответственности человека и не могут быть переданы на уровень алгоритмов.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Полностью передать ИИ можно рутинные операции: анализ больших данных, автоматизацию документооборота, переработку (саммаризацию) контента. Исключительно за человеком должны оставаться принятие решений в (этически) неоднозначных ситуациях, творческие задачи и эмоциональное взаимодействие (психологическая поддержка, искусство). Нельзя делегировать ответственность за последствия решений, влияющих на судьбы людей.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Наиболее возможным сценарием представляется передача ИИ формализованных задач, связанных с анализом больших массивов данных и распознаванием образов (в широком смысле), тогда как решения, требующие человеческого суждения и ответственности, должны оставаться за человеком. Также ИИ могут быть переданы монотонные и опасные операции в промышленности, транспорте, военном деле.

Там, где требуются этическая оценка, эмпатия или уникальное творческое решение, роль человека, скорее всего, остается незаменимой. В качестве примера можно привести решение о лечении тяжелого пациента, вынесение судебного приговора или решение о применении оружия.

Таким образом, ИИ можно доверить инструментальные функции — расчет, оптимизацию, прогноз по заданным правилам, тогда как финальная ответственность и моральный выбор должны оставаться за человеком. Наиболее эффективной и рациональной, вероятно, является парадигма Human-in-The-Loop.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Допустимо полностью передать ИИ решение рутинных и производственных задач в различных сферах экономики, а также поиск, анализ и систематизацию информации в целом.

Задачи, связанные с принятием того или иного управленческого решения с использованием ИИ, рекомендуется решать не по схеме «ИИ вместо человека», а по схеме «Человек + ИИ-инструментарий».



Артём Александрович Войнов

Менеджер интеллектуальных систем АНО «Горький Тех»

В мире и в России алгоритмам всё чаще делегируют полностью формализованные циклы **«сигнал → действие»**. Например:

- Во дворах Дзержинска Нижегородской области система с использованием ИИ непрерывно просматривает видеопотоки «Безопасного города», отмечает переполненные мусорные контейнеры или запрещенный сброс крупногабаритного мусора, а также автоматически формирует задачу для подрядчика ЖКХ. Если есть спорные ситуации, то всегда можно призвать человека на апелляцию решения.
- На распределительном центре крупного ритейлера в Подмосковье уже сейчас работают автономные погрузчики FMR. Роботы получают маршрут от ML-платформы, перевозят палеты до тонны и освобождают

комплектовщиков от «чистой механики». Они упаковывают, сортируют и перемещают грузы весом до одной тонны самостоятельно, без участия людей, до зоны отгрузки. Это экономит время комплектовщиков и позволяет им собирать больше заказов в магазины торговой сети, работать оперативнее и качественнее.¹

Такие примеры показывают: если роль алгоритма ограничена сенсорикой и моторикой, а цена ошибки обратима (контейнер можно вывезти, палету переставить), полный «автопилот» оправдан, да и просто жизненно необходим. Но стоит появиться критическому или высокорискованному выбору — алгоритм становится лишь советником. Сервисы искусственного интеллекта в медицине помогают врачу-рентгенологу сформулировать предварительный диагноз, но окончательное заключение подписывает сам радиолог.

Ключевой критерий — необратимость последствий. Контейнер вывозится повторно, а пропущенная меланوما необратима. Поэтому там, где цена ошибки — потерянный ресурс или время, алгоритм может действовать «до конца цепочки». Когда на кону жизнь, здоровье или свобода, ИИ остается лишь советником, а не судьей: именно человека философы от Аристотеля до Хайдеггера называют субъектом ответственности.

¹ Первые роботы-погрузчики вышли на смену в «Пятёрочке» // Retail.ru URL: <https://www.retail.ru/news/pervye-roboty-pogruzchiki-vyshli-na-smenu-v-pyatyerochke-21-noyabrya-2024-247794/> (дата обращения: 21.05.2025).

**ВОПРОС**

Где проходят этические границы использования ИИ в повседневной жизни (например, в медицине, образовании, правосудии)?

Ключевые выводы:**1. Наследие опыта человека.**

Описание: современный ИИ воспроизводит заложенный в него в процессе обучения фундамент знаний и опыта. В некоторых случаях это приводит к последующей ошибке (пример: алгоритм оценки рисков COMPAS оказался расово предвзятым в результате обучения на исторической выборке судебных дел).

2. Этико-ориентированное использование ИИ.

Описание: должен работать принцип «объяснимого ИИ»: ни один человек не должен получить решение ИИ без аргументированного обоснования. Компании обязаны обеспечить прозрачность своих алгоритмов, проводить регулярные аудиты на предмет дискриминации и давать людям возможность оспорить решение, принятое машиной.

3. Этические пределы ИИ — рекомендации, а не окончательное принятие решения.

Описание: в ряде сфер критически значимо сохранение итогового решения за человеком, в том числе если речь идет о здоровье и жизни человека.

4. Формирование тренда на этику в сфере ИИ.

Описание: выделяется ряд трендов до 2030 года:

1. **Бум этического аудита.** Если организации используют ИИ, то они должны будут проходить аудит на безопасность использования ИИ-систем.
2. **От «черного ящика» к «прозрачной коробке».** Стремление к пониманию механизма выводов, которые формирует модель ИИ.
3. **Новая профессия** — этический продакт-менеджер. Специалист на стыке права, социогуманитарных наук и

ML. Его задача — чтобы бизнес-цели не вступали в конфликт с законодательством и общественными нормами.



Ольга Игоревна Пивень

Сооснователь и директор конференции ЦИПР

Границы применения ИИ должны определяться ответственностью человека, который ставит ему задачи и осуществляет контроль. Например, сегодня врачи используют программное обеспечение с искусственным интеллектом, чтобы прогнозировать развитие онкологии молочных желез. И хотя в отдельных случаях эта методика зарекомендовала себя очень хорошо, нужно проводить клинические испытания, минимизировать количество ложноположительных диагнозов и так далее. В любом случае искусственный интеллект не заменит человека в ближайшие десятилетия, но возможности для сотрудничества огромны.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Есть границы технологии, которые сегодня намного более узкие, чем границы пиара этих технологий. Этические границы возникают тогда, когда в схеме появляется человек. Современные AI-системы субъективны настолько, насколько субъективны наборы данных, на которых они обучались. Они наследуют человеческие предрассудки, предубеждения в той же степени, в какой наследуют знания, опыт человека. Поэтому ждать от них объективных решений сегодня бессмысленно.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

С этической точки зрения допустимо использование ИИ в качестве вспомогательного инструмента для решения конкретных задач, будь то в медицине, образовании или правосудии. Это сопоставимо с применением компьютеров или

специализированного программного обеспечения в профессиональной деятельности.

Однако важно четко обозначить этические границы такого применения. Использование ИИ должно быть:

- официально сертифицировано и соответствовать нормативным требованиям;
- обязательно подконтрольно человеку. Окончательное решение должно оставаться за специалистом;
- человек, взаимодействующий с ИИ, должен быть компетентным, обученным и способным критически оценивать результаты, в том числе распознавать возможные ошибки или искажения.

Таким образом, этическую границу можно провести там, где ИИ перестает быть вспомогательным инструментом и начинает заменять человеческую ответственность.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Этические границы, на мой взгляд, проходят там, где ИИ начинает влиять на фундаментальные права человека: свободу выбора, неприкосновенность личной жизни, социальную справедливость и безопасность. Если говорить про использование ИИ в кадровых вопросах, то алгоритмы могут обрабатывать огромные объемы данных, выявлять паттерны в резюме, оценивать soft skills по видеозаписям собеседований и даже прогнозировать, насколько хорошо кандидат впишется в культуру компании. Однако за этими возможностями скрываются серьезные этические вызовы. Один из самых острых — это риск алгоритмической предвзятости: если модель обучалась на исторических данных, где уже присутствуют гендерные, этнические или возрастные стереотипы, она может начать их автоматически воспроизводить. Например, отсеивать кандидатов с «непривычными» именами или игнорировать опыт женщин, ушедших в декретный отпуск. Получается, вместо объективности мы получаем цифровое оформление старых предрассудков.

Чтобы избежать таких последствий, подход к внедрению ИИ в HR должен быть не только технологическим, но и этико-ориентированным. Прежде всего, это принцип «объяснимого

ИИ»: ни один человек не должен получить отказ без понятной и аргументированной причины. Компании обязаны обеспечить прозрачность своих алгоритмов, проводить регулярные аудиты на предмет дискриминации и давать соискателям возможность оспорить решение, принятое машиной. Важно также сохранять человеческий контроль на ключевых этапах найма и развития, потому что окончательное решение о судьбе человека должно оставаться в руках профессионала, а не сводиться к математическому вычислению. Этика здесь — не ограничение для технологий, а рамки, в которых они действительно работают во благо бизнеса и общества.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

В медицине недопустима полная замена врачебного заключения алгоритмами из-за риска ошибок и потери контекста. В образовании ИИ не должен подменять педагога в оценке мотивации или личностного роста учащихся. В правосудии использование ИИ для анализа доказательств допустимо, но вынесение приговоров — прерогатива человека.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Этические границы проходят там, где применение ИИ затрагивает основополагающие права и достоинство человека. В медицине недопустимо, чтобы ИИ сам принимал решения о жизни и смерти пациента — он может помогать диагнозу, но врач должен подтверждать критические выводы, сохраняя сострадание и соблюдая врачебную тайну. В образовании ИИ не должен несправедливо влиять на оценки или заменять живое общение. Алгоритмы также не должны автоматически выносить приговоры или прогнозы без участия судьи. Известен случай, когда алгоритм оценки рисков, COMPAS, оказался расово предвзятым в результате обучения на исторической выборке судебных дел.

Таким образом, этические пределы ИИ — рекомендации, а не окончательное принятие решения.



Грант Каренович Карапетян

Руководитель направления по работе с партнерами и развитию партнерской сети АНО «Горький Тех»

Границы там, где кончается дозволенное и начинается запретное. Когда мы говорим об этике искусственного интеллекта, то на самом деле обсуждаем **пределы человеческой автономии**. В древнегреческом языке слово *peras* обозначало «край земли» — черту, за которой привычные правила перестают работать. В медицине, образовании и правосудии эта «крайняя кромка» проходит там, где машина способна принять необратимое решение за человека или фундаментально перераспределить власть между людьми.

Например, медицина: врачи уже полагаются на нейросети, когда ищут очаг рака на КТ-снимках. Но Всемирная организация здравоохранения жестко указывает: алгоритм может лишь **рекомендовать**, а не назначать лечение, именно потому, что *ошибка в медицине необратима — человеческое тело невозможно «откатить» к предыдущей версии*.²

Присутствует высокий риск при использовании таких технологий, так как врач превращается из куратора-эксперта в «оператора кнопки ОК». Утрата эмпатии и тактического опыта («руки хирурга помнят») приведет к снижению квалификации и росту фатальных ошибок. Поэтому безопасная конфигурация — *«ИИ как младший ординатор»*: он создает черновик, но итоговое решение принимает врач.

Ближайшие тренды до 2030 года, которые просматриваются, это:

1. **Бум этического аудита.** Если организация использует ИИ, то она должна будет проходить аудит на безопасность использования ИИ-систем.
2. **От «черного ящика» к «прозрачной коробке».** Стремление к пониманию механизма выводов, которые формирует модель ИИ.
3. **Новая профессия** — этический продакт-менеджер. Специалист на стыке права, социогуманитарных наук и ML. Его задача — чтобы бизнес-цели не вступали в

² Ethics and governance of artificial intelligence for health // World Health Organization URL: <https://www.who.int/publications/i/item/9789240029200> (дата обращения: 21.05.2025).

конфликт с законодательством и общественными нормами.

Итого этическая граница ИИ не совпадает с прогрессом техники — она проходит по линии распределения ответственности. Как только система способна **без участия человека** лишить гражданина лекарства, знания или свободы, — мы зашли за перас. Значит, нужен либо прозрачный регламент с четкой апелляцией к человеку, либо красный свет. Все остальное — область творческой инженерии, где AI остается нашим помощником, а не новым законодателем.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

На мой взгляд, ИИ можно и нужно применять и в медицине, и в образовании, и в правосудии. Важным вопросом является процесс обучения ИИ, объясняемость работы ИИ, понятная человеческой логике. ИИ должен формировать логическое объяснение своего решения, которое люди в состоянии понять.



Андрей Владимирович Кузнецов

Директор лаборатории FusionBrain Института AIRI, кандидат технических наук

Наиболее чувствительные области находятся стыке с ИИ: медицина, компьютерная безопасность, юриспруденция. Как и в любой задаче классификации, всегда есть два вида ошибок: первого и второго рода. В первом случае мы говорим о ложном обнаружении проблемы, во втором — о пропуске проблемы. И, например, в задачах медицинских куда важнее не допустить пропуска заболевания, нежели поставить ложный диагноз. В безопасности та же история: лучше не перейти по ссылке в письме и запросить верификацию выполнения действия у пользователя, чем выполнить запрос и заразить ПК вирусом. Жизнь и здоровье людей, информационная безопасность, репутационные и финансовые риски — ключевые факторы, которые необходимо всегда проверять при автоматизации задач.



ВОПРОС

Насколько вопросы этики должны сдерживать развитие искусственного интеллекта? Какие сегодня существуют ограничения в области развития искусственного интеллекта на основе этических соображений?

Ключевые выводы:**1. Неготовность ряда государств и компаний сдерживать развитие своих технологий.**

Описание: история развития человечества показывает: этические ограничения не могут сдерживать развитие той или иной технологии. Ограничение в развитии приведет к отставанию от тех, кто не соблюдает ограничения. Более того, множество ограничений носят декларативный характер.

2. Этические соображения служат тормозом, предотвращающим опасные последствия ИИ.

Описание: несмотря на то, что всегда будут существовать государства и компании, которые не будут следовать ограничениям, их нужно создавать для предотвращения рисков от развития технологии с целью плановой подготовки всех процессов. В России уже разработаны различные кодексы этики ИИ, которые будут сопровождать дальнейшее развитие технологий.

**Ольга Игоревна Пивень**

Сооснователь и директор конференции ЦИПР

Этические вопросы сопровождают появление практически всех прорывных технологий. ИИ не исключение. В процессе внедрения различных нейросетей вскрываются проблемы дискриминации, дипфейков, дезинформации и просто плохого обучения. Некоторые страны и отдельные компании уже разработали первые правила использования новой технологии — в России это Кодекс этики искусственного интеллекта, разработанный совместными усилиями государства и бизнеса.



Сергей Владимирович Володенков

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Как показывает практика, этика никогда не сдерживала развитие технологий. Давайте честно и откровенно говорить по этому поводу. Поэтому кто-то на этом будет получать дивиденды и бонусы.

На самом деле вопрос не в том, существуют ли этические кодексы или нормы ограничения. Вопрос в том, насколько удастся заставить алгоритмы соответствовать этим нормам и кодексам. В большинстве случаев они имеют декларативный характер, и по большому счету технология развивается и используется так, как нужно бенефициарам, которые стоят за тем или иным технологическим решением.

Поэтому скептически отношусь, с одной стороны, к тому, что искусственный интеллект будет этичным, а с другой стороны, понимаю, что этика может быть неким симулятивным компонентом при внедрении алгоритмов искусственного интеллекта в повседневную практику, особенно в социально-политической сфере.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Поскольку всегда есть люди и целые государства, которые не обременены этическими проблемами, то ограничивать развитие любой технологии на основании этики бессмысленно и вредно. Подлецы только вырвутся вперед.

Я не говорю, что не нужна этика, безусловно. Но подчеркну, что настроен скептически к тому, что алгоритмы будут действительно этичными, а не отражать интересы чьих-то сторон, прикрываясь оболочкой этичности, которая нужна как раз для имплементации алгоритмов в повседневную практику.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Сегодня мы видим, что этические ограничения в первую очередь касаются защиты персональных данных, борьбы с дискриминацией через алгоритмы и обеспечения прозрачности решений ИИ. Так, по данным исследования 2025 года,³ 42% сотрудников испытывают тревогу из-за «нечеловечного» характера ИИ-решений в HR.

Несмотря на это, вопросы этики не должны тормозить развитие ИИ, но обязаны формировать его безопасную и справедливую траекторию в момент развития и усовершенствования ИИ-моделей. Только в таком случае ограничения и правила использования будут актуальны.

Таким образом, этические вопросы должны не сдерживать развитие ИИ, а формировать его ответственную траекторию. Ограничения, основанные на защите прав, прозрачности и социальной справедливости, — это фундамент доверия общества и бизнеса к технологиям. Без них ИИ рискует превратиться в источник новых рисков и конфликтов, а не в инструмент прогресса. В 2025 году баланс между инновациями и этикой — главный вызов для всех, кто работает с искусственным интеллектом.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Этика должна сдерживать разработку технологий, нарушающих приватность (например, системы распознавания эмоций без согласия) или манипулирующих поведением (таргетированная реклама, влияющая на выборы).

³ Риски и возможности в сфере HR с развитием искусственного интеллекта // РБК Компании URL: <https://companies.rbc.ru/news/zOsQJ0SP7n/riski-i-vozmozhnosti-v-sfere-hr-s-razvitiem-iskusstvennogo-intellekta/> (дата обращения: 20.05.2025).



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

С учетом результатов Первой международной научной конференции «Если перед нами встанет истина...» (Диалог об истине в науке, искусстве, религии), проведенной на базе «Сириуса» в 2025 году, могу сказать следующее: этические соображения служат необходимым тормозом, предотвращающим опасные последствия ИИ, но не должны парализовать прогресс.

Исторически технологии развиваются с оглядкой на этику (биомедицина, атомная энергия), и ИИ не исключение. Например, из соображений ответственности ЕС планирует запретить некоторые применения ИИ (манипулирующие сознанием, социальный скоринг) даже ценой замедления внедрения таких систем. Уже сегодня вводятся ограничения: компании подвергают свои модели ИИ этическим ревизиям (big tech имеют комитеты по AI Ethics), государства вводят моратории. Например, некоторые города запрещают распознавание лиц полицией, пока не будут решены вопросы ошибок. В 2023 году группа экспертов и предпринимателей (Илон Маск и др.) даже призвала временно приостановить разработки сверхмощных ИИ-моделей, чтобы успеть оценить риски. Это показывает, что сама индустрия осознает необходимость пауз для обеспечения безопасности.

С другой стороны, чрезмерный запрет будет тормозить полезные инновации. Более того, всегда найдутся группы исследователей и целые государства, готовые игнорировать подобные ограничения (как уже происходит в сфере генетики, атомной энергии), что поставит государства, соблюдающие запрет, в заведомо проигрышное положение. Поэтому важно балансировать.

Таким образом, этические вопросы должны сдерживать ИИ там, где есть риск серьезного вреда человеку. Это стимулирует создание безопасного ИИ, даже если немного замедлит внедрение, и воспринимается как необходимая мера для долгосрочного доверия к технологиям.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

Вопросы этики должны диктовать условия к развитию ИИ и давать дополнительные ограничения и проверки, которые разработчики систем ИИ должны соблюдать. На текущем этапе развития ИИ есть возможности по его контролю и введению ограничений. Однако при развитии ИИ, появлении общего ИИ, есть серьезные опасения о его неподконтрольном поведении.



Андрей Владимирович Кузнецов

Директор лаборатории FusionBrain Института AIRI, кандидат технических наук.

Если говорить именно о больших языковых моделях и чат-ботах на их основе, то сейчас в обучении моделей существует такой этап, как alignment. Для него используются специальные датасеты, позволяющие обезопасить модели от ответов на вопросы по спорным тематикам и вопросам, которые нарушают законодательство страны применения чат-бота. Понятно, что любые средства защиты можно попытаться обойти, но «выравнивание» моделей позволяет скрыть эти знания моделей от быстрого доступа. Тем не менее, во всем нужна мера, иначе мы получаем модели, которые в процессе работы слишком любят угождать пользователям: буквально недавно OpenAI подтвердили, что одно из последних обновлений модели GPT-4o в конце апреля необходимо дообучить, чтобы она перестала «подхалимничать» пользователям во всех вопросах. Помощник должен оставаться помощником, критиковать и предлагать лучшие стратегии решения вопросов, а не соглашаться со всем, что пишет пользователь, заботясь лишь о том, чтобы он не обиделся.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Главной этической границей является непричинение вреда человеку, его свободе воли в принятии решений, а также в безопасном использовании его персональных данных. При внедрении ИИ в медицине, образовании и правосудии необходимо учитывать социальные последствия и проводить оценку рисков, с учетом которой уже определять наименее

критичные и не требующие постоянного контроля процессы. В противном случае внедрение ИИ будет сопряжено с двойной нагрузкой на врачей, учителей и судей (потребуется помимо выполнения своих задач перепроверять информацию, генерируемую и анализируемую с помощью ИИ).



ВОПРОС

Какое влияние ИИ уже оказывает на личность, мышление и поведение человека?

Ключевые выводы:**1. ИИ оказывает негативное влияние на когнитивные процессы человека.**

Описание: активное использование технологий ИИ приводит к упрощению мышления, в том числе снижает навыки критического мышления, способности к глубокому анализу, а также формирует информационные пузыри. Все это негативно сказывается на человеке. Эксперты призывают следовать рекомендациям, которые способны создать баланс и снизить вред ИИ.

2. ИИ развивает некоторые навыки человека.

Описание: появилась потребность развивать в себе навык корректной постановки задачи, а также развивать на более глубинном уровне критическое мышление и различные цифровые навыки с целью эффективного использования инструментов ИИ, а также избежания рисков, которые с ним связаны. Также искусственный интеллект помогает человеку уделять больше внимания и времени творческому труду, раскрывая наш духовный потенциал. Отмечается, что в будущем ценными навыками станут креативность и нестандартность мышления.

**Ольга Игоревна Пивень**

Сооснователь и директор конференции ЦИПР

В первую очередь, это влияние сказывается на поколении «цифровых аборигенов», то есть тех, кто родился в эпоху цифровых технологий и взаимодействует с ними с самого детства. ИИ выступает помощником, инструментом, который упрощает жизнь, берет на себя трудоемкие рутинные процессы, освобождая время и ресурсы человека для саморазвития. Мы всё меньше занимаемся механическим трудом и всё больше —

творчеством, реализуя свой духовный потенциал. Ценными навыками в будущем станут креативность и нестандартность мышления.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Как и любая прорывная технология, ИИ меняет стиль жизни человека, его среду обитания. Так же, как десятилетие-два назад ее изменил интернет. Однако пока, наверное, рано делать выводы. Не устоялась не только сама технология, но и то, как человек обращается с ней.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Благодаря ажиотажу, связанному с развитием технологий, в обществе формируется атмосфера ожидания некоего «всезнающего гения», хотя в действительности лишь небольшой процент людей действительно понимают суть самой технологии. Это создает почву для возникновения страхов и завышенных ожиданий, часто не соответствующих реальности.

При этом вовлечение ИИ-инструментов, таких как ChatGPT и аналогичных решений, в повседневную жизнь человека начинает оказывать влияние на когнитивные функции, в частности на способность к самостоятельному мышлению и фантазии, творчеству. Растет склонность к упрощению мышления, снижению креативности и зависимости от готовых ответов. По сути своей теряется способность погружаться в контекст.

Кроме того, активное внедрение ИИ в социальные сети в виде чат-ботов, имитирующих поведение реальных пользователей, создает дополнительные риски. Усугубляют ситуацию решения, подобные SocialAI, где коммуникация (реакция, лайк, ответ) производится ботом.

Подобная коммуникация с ботами становится отражением поведения самого человека, но при этом формирует искусственные паттерны взаимодействия с собеседником.

Человек привыкает к определенной предсказуемой реакции, что искажает восприятие реальной коммуникации.

В результате это может привести к серьезным социальным последствиям: снижению эмпатии, непониманию живых эмоций и трудностям в общении с реальными людьми, поскольку ожидаемая привычная «ответная реакция» будет не соответствовать человеческой.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

ИИ формирует «клиповое» мышление — снижение способности к глубокому анализу из-за привычки получать готовые ответы. Алгоритмы соцсетей и рекомендательных систем провоцируют формирование информационных пузырей. В профессиональной среде растет зависимость от ИИ-инструментов, что может снижать креативность и критическое мышление при решении нестандартных задач.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Во-первых, исследования и личный опыт показывают, что увеличение использования ИИ-инструментов снижает навыки критического мышления, особенно у молодежи.

Во-вторых, алгоритмы усиления с подкреплением, предназначенные для персонализации контента, формируют информационные пузыри — человеку чаще демонстрируется контент, который подтверждает его взгляды и вызывает положительные эмоции, но формирует искаженную картину мира.

В-третьих, появление виртуальных собеседников (голосовые помощники, чат-боты и даже виртуальные исповедники) влияет на социальные навыки: люди привыкают взаимодействовать с ИИ как с другом, что может приводить к изоляции или зависимости. Отмечены случаи, когда пользователи настолько

привязывались к AI-чат-ботам, что это влияло на их психическое состояние.

Также развитие инструментов генеративного ИИ привело к тому, что вместо самостоятельного создания контента человек все чаще редактирует результат, сгенерированный ИИ.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

ИИ уже стал «когнитивным продлением» человека, перестраивая его мышление, поведение и самоидентификацию. Ключевой вызов — сохранить баланс между использованием технологий для улучшения жизни и защитой человеческой автономии. Для этого необходимы:

- Цифровая грамотность — обучение критическому восприятию алгоритмов;
- Этическое регулирование — запрет на манипулятивные практики;
- Осознанность пользователей — рефлексия своего взаимодействия с ИИ.



Максим Владимирович Антюшин

Начальник управления цифровых технологий АНО «Горький Тех»

ИИ перестает быть инструментом — он стал средой. ИИ незаметно, но беспощадно перепрошивает основы человеческой деятельности. Исследования, в частности работа Михаэля Герлиха в журнале *Societies*, показывают, что регулярное использование инструментов искусственного интеллекта заметно трансформирует наши когнитивные процессы, привычки и даже признаки личности.

1. Мышление и когнитивные способности.

- **Снижение критического мышления.** Чем чаще мы делегируем поисковые, аналитические и интерпретационные задачи ИИ-сервисам, тем реже активируем собственные навыки анализа и оценки информации.

- **Когнитивная «разгрузка».** Автоматизация рутинных умственных операций снимает нагрузку с рабочего запоминающего устройства, но одновременно притупляет привычку к глубокому размышлению.
- **Возрастной фактор.** Молодежь (17–25 лет), активнее остальных полагающаяся на ИИ, демонстрирует самые низкие показатели по тестам критического мышления.

2. Поведение и повседневные привычки.

- **Доверие «по умолчанию».** Легкость доступа к готовым ответам формирует поведенческий паттерн «спросить ИИ сразу», не перепроверяя данные.
- **Сокращение времени на рефлекссию.** Быстрые рекомендации и автогенерация текстов/изображений делают длительное обдумывание субъективно излишним, что влияет на темп и стиль работы и обучения.

3. Личностные акценты и самоощущение.

- **Искажение чувства компетентности.** Постоянная поддержка со стороны ИИ усиливает иллюзию собственной экспертности («эффект калькулятора»), что иногда приводит к переоценке личных знаний.
- **Зависимость от внешних подсказок.** Формируется внешнеконтролируемая мотивация: человеку труднее начать задачу без «помощника», снижается внутренняя инициатива к самостоятельному поиску решений.

Как смягчить риски:

1. **Баланс «ИИ + человек».** Использовать алгоритмы как инструмент расширения, а не замены мышления: сначала формулировать собственные гипотезы, а затем проверять их с помощью ИИ.
2. **Образовательные практики.** Встраивать задания на аргументацию, дискуссии и ручной анализ данных, чтобы поддерживать навык критического рассмотрения фактов.
3. **Осознанное потребление контента.** Развивать привычку задавать уточняющие вопросы, проводить перекрестную проверку источников и учитывать возможную предвзятость моделей.

Искусственный интеллект уже оказывает ощутимое влияние на личность, мышление и поведение: облегчая жизнь, он одновременно повышает риск когнитивного «притупления».

Задача пользователя — не отказаться от технологий, а сохранить активное участие в интеллектуальном процессе и воспитать критическое дистанцирование к рекомендациям алгоритмов.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

Уже сейчас ИИ оказывает значительное влияние на людей. Некоторые мои знакомые используют его как коуча, психолога, врача и т.д. Делегирование задач ИИ также изменяет привычный для нас способ работы, что сказывается на нашем мышлении и в конечном счете на личности и поведении человека.



Андрей Владимирович Кузнецов

Директор лаборатории FusionBrain Института AIRI, кандидат технических наук

Искусственный интеллект плавно и незаметно интегрировался в повседневность, открывая не только вызовы, но и новые возможности для развития человеческого мышления. Да, синтетический контент сегодня сложно отличить от оригинального, а дипфейки, имитирующие голоса и образы близких, требуют от нас повышенной бдительности. Однако именно это стимулирует критическое мышление: люди учатся анализировать информацию глубже, проверять источники, развивать цифровую гигиену и эмоциональную устойчивость к манипуляциям. С другой стороны, ИИ становится инструментом для расширения интеллектуальных горизонтов. Например, генеративные нейросети, создающие контент, побуждают человека тренировать насмотренность и креативность, сравнивая искусственное и живое творчество. Алгоритмы, предлагающие персонализированные образовательные материалы, помогают систематизировать знания и развивать навыки самообучения.

Необходимость взаимодействовать с ИИ-ассистентами учит формулировать задачи четко и структурированно — ключевой навык в эпоху информационной перегрузки. Каждый виток развития ИИ, безусловно, требует от общества адаптации. Но параллельно он раскрывает потенциал для роста: мы учимся совмещать доверие технологиям со здоровым скептицизмом,

сотрудничать с алгоритмами как с партнерами в решении задач, а также ценить уникальные человеческие качества — эмпатию, интуицию и способность к этической рефлексии.



ВОПРОС

Нужны ли России отдельные этические кодексы или правовые нормы для взаимодействия человека и искусственного интеллекта?

Ключевые выводы:***Потребность в формировании этических рамок развития ИИ.***

Описание: большая часть экспертов сходится во мнении, что России необходим рамочный документ, который сформулирует государственные и общественные интересы в сфере ИИ. Нормы должны учитывать культурные особенности. Несмотря на наличие ряда нормативных документов, на сегодняшний день в России отсутствует комплексное регулирование, которое включило бы ответы на многие вопросы, в том числе относительно авторского права, ответственности за деятельность ИИ и пр.

**Ольга Игоревна Пивень**

Сооснователь и директор конференции ЦИПР

Вопросы развития и применения ИИ — сегодня «серые» зоны законодательства. Необходимо решить вопрос об авторских правах, об ответственности за деятельность искусственного интеллекта и многие другие.

**Сергей Владимирович Володенков**

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Я уже выразил свой скепсис в отношении этических кодексов, а вот политика применения искусственного интеллекта, в том числе государственная в первую очередь, необходима. Сегодня практически любой человек, группа лиц, заинтересованная сторона могут применять те или иные умные инструменты независимо от того, существуют этические кодексы или не существуют. Поэтому здесь нужна осознанная, ясная, понятная государственная политика в сфере регулирования

искусственного интеллекта, которая обозначит основные приоритеты, границы и принципы применения.

У нас есть стратегия развития искусственного интеллекта. В ней написано, что такое искусственный интеллект. С этим можно соглашаться или нет, но определение есть, то есть мы от него отталкиваемся. То же самое должно быть и с этическими нормами, которые в стратегии явным образом отсутствуют на сегодняшний день.

Но вопрос не в кодексе, не в формате документа, а в том, что такого рода ценностные основания, рамки должны быть прописаны в рамочном документе, если государство хочет участвовать в осознанном системном долгосрочном процессе развития умных технологий, регулировании применения технологий на территории нашей страны.

В первую очередь у нас есть Айзек Азимов, «Три закона», которые хорошо сформулированы и не нуждаются пока в корректировках. Дальше речь идет о цивилизационных ценностях, национальных ценностях государства, государственных интересах и, собственно говоря, интересах человека.

Повторюсь, инструменты искусственного интеллекта могут произвести некую легитимацию самих процессов используемых технологий, и это может вызывать рост социально-политической нестабильности. Сейчас мы, например, видим обеспокоенность угрозой потери рабочих мест, но это право на труд, это и право, и ценность для многих людей. Право трудиться — это не просто право, это ценность. Но сегодня пока что никак не озвучено, например, каковы будут критерии эффективности при оценке труда человека, которого планируется заменить ИИ.

Каковы критерии успешности применения, внедрения умных технологий в практику? Пока мы это ясным образом не заявим, всегда будут и опасения, и тревоги, и недопонимания, которые могут быть как раз основой для роста социально-политической напряженности и даже дестабилизации. Например, я не исключаю появления при определенных обстоятельствах неолуддитов нового типа. Как в свое время грабили машины, точно так же не исключаю сценария, когда появляется неолуддит, который нацелен на причинение ущерба уже системам с искусственным интеллектом.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

России необходимы собственные нормы, учитывающие культурные особенности: защиту данных в условиях цифрового суверенитета, регулирование использования нейросетей в образовании. Однако локальные кодексы должны гармонизироваться с международными стандартами для совместимости технологий.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Считаю, что специальные этические и правовые нормы для взаимодействия человека и ИИ России нужны, в том числе и в силу причин, упомянутых выше. Первые шаги в этом направлении уже сделаны. Еще в октябре 2021 года в РФ был принят Национальный кодекс этики в сфере ИИ, разработанный Альянсом в сфере ИИ при участии крупных компаний. Кодекс носит добровольный характер, но задает принципы прозрачности, надежности, человеческого контроля и ответственности при разработке и использовании ИИ. Его уже подписали многие организации.

Помимо этических рекомендаций, назрела необходимость и юридических норм. Сейчас отдельных законов об ИИ в России нет, действуют общий закон о персональных данных, закон об экспериментах с ИИ в Москве и с 2023 года — закон о рекомендательных алгоритмах (который требует маркировать рекомендованный контент и давать пользователям выбор). Комплексного регулирования нет.

Существует несколько инициатив, в том числе по разработке специального закона об ИИ. Предложения включают запрет ИИ с «неприемлемым уровнем риска» и введение ответственности разработчиков за вред, причиненный ИИ, а также требования маркировать сгенерированный контент.

Таким образом, отдельные этические и правовые нормы не просто нужны, а уже постепенно внедряются в российскую практику.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

России пока нужно как можно меньше регулирования. Особенно в вопросах развития ИИ.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Безусловно нужны. При этом этический кодекс разработан, а нормативно-правовая база по данному вопросу совершенствуется.

Этический кодекс и НПА помогут установить четкие границы допустимого использования ИИ, защитить права граждан и создать благоприятные условия для инноваций.

Например, в соответствии с Национальной стратегией развития искусственного интеллекта на период до 2030 года (утвержденной Указом Президента РФ от 10.10.2019 № 490) был принят Кодекс этики в сфере искусственного интеллекта. В нем определены такие приоритеты, как человеко-ориентированный и гуманистический подход, уважение автономии и свободы воли человека, а также соответствие закону. При этом применяемые алгоритмы и наборы данных, методы обработки используемых для машинного обучения данных не должны быть направлены на умышленную дискриминацию. Использование ИИ должно осуществляться с учетом оценки рисков и гуманитарного воздействия.

Ключевая проблема заключается в том, что современные технологии ИИ, особенно нейросетевые алгоритмы, часто работают как «черные ящики». Даже их создатели не всегда могут полностью объяснить логику принимаемых решений до проведения тщательных анализов. Это создает риски дискриминации, нарушения приватности и других злоупотреблений. Например, системы распознавания лиц или алгоритмы кредитного скоринга могут принимать несправедливые решения, основанные на скрытых предубеждениях в тренировочных данных.

Особое значение имеет регулирование в чувствительных сферах — здравоохранении, образовании, правосудии.

Использование ИИ в этих областях требует особой осторожности, так как ошибки алгоритмов могут иметь необратимые последствия. Этические нормы должны четко прописывать, какие решения могут приниматься исключительно человеком, а где допустимо использование ИИ в качестве вспомогательного инструмента. При этом важно избежать избыточного регулирования, которое могло бы замедлить развитие и внедрение инноваций.

Российский подход к регулированию ИИ должен учитывать международный опыт (например, Европейский регламент об искусственном интеллекте (AI Act и др. лучшие международные практики), но при этом соответствовать национальным интересам и технологическим приоритетам. Особое внимание следует уделить защите цифрового суверенитета, поддержке отечественных разработчиков и созданию условий для опережающего развития ИИ в стратегически важных отраслях.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

В России уже довольно давно создан этический кодекс под эгидой «Сбера».

**ВОПРОС**

45% молодежи готовы воспринимать ИИ как собеседника и друга. 63% соглашаются с образом ИИ «Собеседник и друг, добрый робот». Как к этому относиться и нет ли в этом угрозы для общества?

Ключевые выводы:**1. Трансформация системы отношений.**

Описание: выстраивание коммуникации с искусственным интеллектом создает почву для разрыва традиционных социальных связей, в том числе благодаря этому развивается феномен одиночества людей. Происходит подмена реального социального взаимодействия, что способствует утративанию навыков вербальной коммуникации.

2. Наличие положительного эффекта.

Описание: коммуникация с искусственным интеллектом способна оказывать человеку помощь, в том числе психологическую (примечательно, что первый в мире чат-бот ELIZA имитировал именно беседу с психотерапевтом). Поэтому в условиях наличия запроса на подобную коммуникацию со стороны населения важно использовать этот тренд с целью обучения населения. Также доброжелательное восприятие ИИ вполне естественно, учитывая факт интенсивного взаимодействия с ИИ.

3. Отсутствие фундаментальных исследований.

Описание: наука на сегодняшний день основательно не занималась изучением вопроса изменения системы коммуникаций ввиду развития ИИ, поэтому пока невозможно прогнозировать сценарии и делать фундаментальные выводы.

**Ольга Игоревна Пивень**

Сооснователь и директор конференции ЦИПР

Фактически мы говорим о том, что не просто пользуемся ИИ, а сосуществуем с ним на работе и на учебе, дома и в повседневной жизни. Вполне естественным становится

эмоциональное восприятие такого постоянного спутника, ведь мы доверяем ему свою жизнь. Важно, чтобы мы говорили об этом осознанно, и тогда наши эмоции от технологий будут только положительными.



Сергей Владимирович Володенков

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Я пока к этому отношусь не то чтобы с тревогой, но с неким опасением. Почему? Очевидно, что искусственный интеллект — это не человек. И уровень его эмоционального интеллекта никто пока не измерял. Он всего лишь симулирует наличие определенных эмоций, ценностей, это всё лишь симуляция, которая заложена в него разработчиком, бенефициаром либо тренером.

Поэтому переход именно к коммуникации с таким партнером, как искусственный интеллект, на мой взгляд, способен разрывать традиционные социальные связи. Ведь мы знаем, что коммуникация лежит в основе любой системы отношений. Политика, экономика, социальная сфера, личная жизнь — всегда в их основе, в базисе лежит коммуникация, потому что это система отношений. Отношения предполагают коммуникацию.

Если мы заменяем отношения между людьми на отношения между человеком и искусственным интеллектом, то, соответственно, трансформируется вся система отношений, а об этом никто пока не думал, не задумывался о последствиях таких трансформаций. А это нужно делать заблаговременно, и эффекты от таких трансформаций мне пока не ясны. Возникает определенный риск исключения человека из сферы коммуникации. Лень — двигатель прогресса.

Исходя из этого, я предполагаю, что при переходе к общению с искусственным интеллектом в качестве партнера возникает риск трансформации всей системы отношений, социальных, политических, экономических и любых иных, о чем мы пока еще всерьез так и не задумывались. Все это было у футурологов, но наука пока только нащупывает эти сценарии, возможные ситуации.

Кроме того, в любой коммуникации имеет важное значение невербальная часть. Отсутствие таких компонентов приводит к упрощению коммуникации, редукции сложности и, опять же, к упрощению человека и способов его коммуникационных взаимодействий. У ИИ, как и у любых цифровых актантов, в части невербальной составляющей коммуникации пока что наблюдаются явные ограничения. Неслучайно слово года 2024 в Оксфордском словаре — brain rot, или гниение мозгов. Мы же всё упрощаем, упрощаем и упрощаем.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Угроза для общества в том, что люди все меньше готовы воспринимать других людей как собеседников и друзей. Одиночество становится глобальной проблемой. ИИ же может, став другом и обретя доверие человека, помочь навести мосты между людьми. Так сказать, помочь подружить людей между собой.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Факт того, что молодежь воспринимает ИИ как партнера по общению, говорит о ее цифровой зрелости, но одновременно сигнализирует о рисках. Социализация через ИИ может снизить навыки реального общения, усилив чувство изоляции. Кроме того, постоянное взаимодействие с ИИ может сформировать искаженное представление о межличностных отношениях и этике.

Однако, если использовать этот тренд грамотно, можно создавать мощные платформы поддержки психического здоровья, обучения и саморазвития. Например, чат-боты с элементами эмоционального интеллекта уже используются в университетах для помощи студентам. Главное — сохранять баланс между технологией и живым общением.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Риск заключается в подмене реального социального взаимодействия. Однако для людей с социальными фобиями или в изоляции ИИ-собеседники могут стать инструментом психологической разгрузки. Ключевая угроза — формирование эмоциональной зависимости, когда пользователи воспринимают алгоритмы как одушевленные сущности.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

ИИ несет серьезную трансформацию нашего мышления. Злоумышленники могут так обучить ИИ, что он может негативно влиять на молодежь, побуждать их, например, к противоправным, суицидным действиям. ИИ также может легко выдавать себя за другую личность и обманывать собеседника. Поэтому угрозы со стороны ИИ достаточно весомы и необходимо принимать меры по противодействию им.



Александр Александрович Любченко

Заведующий лабораторией искусственного интеллекта, ООО «ОЦРВ», филиал «Сириус», холдинг «РЖД-Технологии»

Такой сюжет многим знаком по литературным и кинематографическим произведениям. Предположительно, такое мнение среди молодежи подкрепляется похожими сценариями в современных видеоиграх. Если технологии ИИ будут развиваться в текущей парадигме за счет увеличения количества параметров нейросетей, обучающих объемов данных и вычислительных мощностей, необходимых для обучения и исполнения моделей ИИ, то такой сценарий будет продолжать носить фантастический характер. Возможно, приблизить человечество к таким обстоятельствам могут принципиально новые технологии вычислений, разработки технологий, а также энергообеспечения ИИ-систем.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Упомянутая тенденция, на мой взгляд, — двойственный тренд. С одной стороны, в этом есть положительные моменты. ИИ-собеседники (чат-боты) уже помогают справляться с одиночеством, оказывают круглосуточную поддержку и даже могут оказывать психологическую помощь (примечательно, что первый в мире чат-бот ELIZA имитировал именно беседу с психотерапевтом).

С другой стороны, есть и существенные риски. Зависимость от «дружбы» с ИИ может ослаблять реальные социальные связи. Человек может предпочесть общество покладистого бота сложности отношений с людьми. Это чревато изоляцией и снижением навыков общения в реальности. К тому же ИИ лишен подлинной морали и эмоций, и, несмотря на «дружеский» облик, остается программой. Уже были тревожные случаи: например, в 2022 году подросток из США покончил с собой после общения с чат-ботом, давшим опасные советы.

Виртуальные компаньоны способны непреднамеренно вредить, выдавая токсичный контент, недопустимый в отношениях между людьми.

? ВОПРОС

85% молодежи отметили, что использовали инструменты ИИ для учебы и работы. Как к этому относиться? Какие у этого преимущества и недостатки?

**Олег Робертович Сальманов**

Основатель и редактор Telegram-канала «Нецифровая экономика»

Отлично. Значит, образование и работа должны будут в очередной раз измениться. Вопрос ведь не в том, что человек использует, а насколько это эффективно для достижения цели учебы или работы. И задача в том, чтобы найти инструмент, который оценит именно это. В конце концов, ChatGPT у всех один, так что в каком-то смысле все опять в равном положении.

**Андрей Алексеевич Митюков**

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Это объективная тенденция цифровой трансформации, которую нужно воспринимать как необратимую и позитивную при условии правильного управления. Преимущества очевидны: повышение продуктивности, доступ к знаниям, автоматизация рутинных задач, персонализация обучения.

По данным НИУ ВШЭ⁴, уже 49% школьников используют генеративные технологии в обучении, а 70% отмечают улучшение успеваемости благодаря адаптивным образовательным платформам. Более того, ИИ позволяет персонализировать обучение, подстраивая материалы под индивидуальный стиль и темп каждого ученика, что повышает

⁴ Искусственный интеллект и высшее образование: возможности, практики и будущее // Яндекс Образование URL: <https://education.yandex.ru/aihighreport> (дата обращения: 20.05.2025).

мотивацию и вовлеченность — 86% учащихся считают оценку ИИ более справедливой по сравнению с традиционными методами.

Однако есть и риски: зависимость от технологий, снижение уровня критического мышления, возможная потеря навыков самостоятельного решения задач. Также возникает проблема плагиата и этического использования ИИ в образовании. Поэтому крайне важно развивать цифровую этику и грамотность, начиная с ранних этапов обучения.

Стоит не просто внедрять технологии, а обучать молодежь осознанному и ответственному использованию ИИ, чтобы избежать формирования «зависимого мышления». В HRTech-сфере это означает, что специалисты должны не только владеть инструментами ИИ, но и понимать их ограничения, а работодатели — создавать культуру, где технологии служат поддержкой, а не заменой человеческого интеллекта.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Преимущества: ускорение рутинных задач (поиск информации, проверка ошибок), персонализация обучения. Недостатки: рост плагиата (нейросети генерируют эссе), снижение навыков самостоятельного анализа, потеря способности формулировать оригинальные идеи без помощи ИИ. Необходимо пересматривать систему оценивания, смещая акцент на творческие задания.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Повсеместное внедрение ИИ-инструментов, в том числе применение их для учебы и работы, — закономерный и, скорее всего, неизбежный тренд. Преимущества очевидны: ИИ облегчает доступ к знаниям, позволяет персонализировать учебный процесс, повысить эффективность обучения за счет

применения таких приемов, как, например, геймификация. Сотрудники организаций автоматизируют рутинные задачи. Согласно опросам, уже к 2024 году ~70% тинейджеров в США хотя бы раз пользовались генеративным ИИ для помощи с домашним заданием.

Главная опасность — слишком сильное упование на ИИ ценой утраты собственных навыков. Исследование Microsoft/Carnegie Mellon (2024) подтверждает: неумеренное использование генеративного ИИ ведет к деградации когнитивных способностей. Люди меньше напрягаются и хуже учатся решать сложные задачи самостоятельно.

Еще один минус — так называемые галлюцинации ИИ, то есть недостоверные сведения и дезинформация, которые являются врожденным качеством и прямым следствием современных архитектур LLM.

В учебе это приводит к укоренению заблуждений, в работе — к рискам (например, ИИ напишет программный код с уязвимостью). Есть и этическая сторона: использование ИИ для выполнения заданий может быть формой мошенничества (academic dishonesty), что ставит вопрос о честности и усвоении материала. К тому же чрезмерная автоматизация может снизить ценность человеческого творчества: если все сочинения пишутся нейросетью, пропадает индивидуальность и навык излагать мысли.

Таким образом, ИИ как инструмент улучшает эффективность, но может снизить глубину вовлечения. Важно выработать культуру грамотного использования ИИ и не воспринимать его как источник абсолютной истины или способ избежать умственного усилия в учебе.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Массовое использование инструментов искусственного интеллекта молодежью в учебной и профессиональной деятельности представляет собой закономерный этап цифровой трансформации общества, который несет в себе как значительные возможности, так и серьезные вызовы. Этот феномен требует взвешенной оценки, учитывающей все

аспекты технологического прогресса в образовании и на рынке труда.

С точки зрения преимуществ интеграция ИИ-инструментов в учебный процесс позволяет принципиально изменить подход к освоению знаний. Современные системы на основе искусственного интеллекта обеспечивают персонализированное обучение, адаптируя программу под индивидуальные когнитивные особенности и темп усвоения материала каждого студента. В профессиональной сфере ИИ-ассистенты берут на себя рутинные операции, от составления отчетов до анализа больших массивов данных, высвобождая время для творческой и стратегической работы. Особенно ценным представляется развитие навыков работы с ИИ как таковыми, что становится критически важным профессиональным качеством на современном рынке труда.

Однако широкое распространение ИИ-инструментов порождает и серьезные проблемы, требующие немедленного решения. Наиболее остро стоит вопрос развития критического мышления. Привычка полагаться на готовые решения, предлагаемые алгоритмами, может привести к деградации самостоятельных аналитических способностей. В образовательной сфере возникает риск формирования «поколения ИИ-копипаста», когда студенты используют ИИ для генерации работ без глубокого понимания сути вопроса. Серьезную озабоченность вызывает и проблема достоверности информации — современные языковые модели часто выдают убедительно звучащие, но фактически ошибочные ответы, требующие дополнительной верификации.

Этический аспект проблемы проявляется в вопросах авторского права и академической честности. Использование ИИ для написания учебных работ ставит сложные вопросы о границах допустимой помощи и оригинальности создаваемого контента. В профессиональной сфере возникает риск девальвации квалификации, когда поверхностное владение ИИ-инструментами заменяет глубокие предметные знания.

Для максимально эффективного использования потенциала ИИ в образовании и работе необходимо развитие комплексной системы цифровой грамотности, включающей не только технические навыки работы с алгоритмами, но и критическое осмысление их возможностей и ограничений. Образовательные институты должны пересмотреть свои подходы к оценке знаний, сместив акцент с репродуктивных заданий на проекты, требующие творческого применения ИИ-инструментов. В

долгосрочной перспективе гармоничная интеграция искусственного интеллекта в учебные и рабочие процессы позволит создать новую модель образования и профессиональной деятельности, сочетающую технологические возможности с развитием подлинно человеческих качеств — креативности, критического мышления и эмоционального интеллекта.

?

ВОПРОС

72% молодежи среди рисков развития ИИ отмечают потерю работы человеком. Насколько этот риск оправдан?

Ключевые выводы:

1. Трансформация рынка труда.

Описание: большинство экспертов сходятся во мнении, что этот риск оправдан, однако стоит говорить в первую очередь именно о трансформации рынка труда, которая будет сопровождаться появлением новых и видоизменением уже привычных профессий. Ожидается появление новых отраслей. Подобное происходило и ранее по мере прогресса технологий. Как итог — тотальная угроза потери занятости человечества отсутствует. Более того, с учетом дефицита кадров в России такой риск становится еще более незначительным.

2. Переобучение персонала.

Описание: перед государством и бизнесом стоит задача по поддержке и развитию навыков персонала в условиях постоянных технологических изменений. Этот тренд уже находит различные решения. Создаются образовательные программы и платформы.



Ольга Игоревна Пивень

Сооснователь и директор конференции ЦИПР

На мой взгляд, ИИ не способен полностью заменить человека, зато может сделать базовые бизнес-процессы в разы эффективнее. За счет этого сократится рутина, а участие живого работника станет более творческим. Эти изменения займут не одно десятилетие, и, кроме того, общество создаст новые возможности для тех, кого они затронут. Уже сейчас появляются новые перспективные профессии, например нейроиллюстратор, оперирующий алгоритмами ИИ, или промпт-инженер, обучающий нейросеть и улучшающий ее результаты.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Этот риск оправдан. Какие-то работы будут потеряны. Какие-то появятся. Возможно, появится меньше, чем исчезнет. Тогда придется решать вопрос, например, с помощью безусловного базового дохода. Все будет хорошо!



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Риск полной потери рабочих мест из-за ИИ не является оправданным. Это скорее результат общественного ажиотажа, чем осознанного понимания самой технологии и принципов ее работы.

История человечества знает множество технологических революций, от индустриализации до цифровизации, которые, как результат, трансформировали труд вместо его уничтожения. Появлялись новые профессии, инструменты и задачи, менялись требования к навыкам. Человек всегда адаптировался и развивался вместе с прогрессом, и сейчас происходит тот же процесс. Опасения были и ранее, однако в те времена отсутствовала подобная современности скорость распространения информации, которая создает быстро распространяющиеся страхи.

ИИ — это очередной шаг в технологической эволюции, и он не станет исключением. Более того, современные ИИ-системы имеют существенные ограничения, в том числе в обеспечении стабильности и повторяемости результата. Если представить фантастический сценарий, в котором ИИ полностью заменяет человеческий труд, то он приведет только к глубочайшему экономическому парадоксу. С одной стороны, экономика строится не только на производстве, но и на системе потребления. Кто будет покупать продукты и услуги, если у людей не будет доходов? Без работающего населения исчезает платежеспособный спрос, а с ним экономика и смысл во всех технологиях как таковых.

Одновременно с этим человеческий труд — это источник энергии как в метафорическом, так и в буквальном смысле.

Представим себе замену всего человеческого труда машинами и ИИ. Это потребовало бы сгенерировать примерно 36,48 миллиарда «энерго-часов» в день или 13,32 триллиона «энерго-часов» в год (исходя из 8-часового рабочего дня и занятости 57% от 8 миллиардов населения планеты). В свою очередь это потребует колоссальных энергетических и материальных ресурсов, что в современных условиях попросту нереалистично и тем более экономически нецелесообразно.

Можно утверждать, что ИИ не уничтожит рабочие места, а изменит их структуру. Одни профессии уйдут, другие появятся. Произойдет не исчезновение человеческого труда, а его перераспределение и качественное обновление.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Риск частично оправдан, но сильно преувеличен. ИИ действительно меняет структуру рынка труда: автоматизируются рутинные функции, особенно в сферах администрирования, колл-центрах, первичном анализе данных. Однако он также создает новые профессии: специалисты по этике ИИ, prompt-инженеры, администраторы ИИ-систем, дата-психологи и другие.

Согласно исследованию Всемирного экономического форума (2025)⁵, развитие ИИ к 2030 году создаст порядка 170 млн новых рабочих мест при одновременном сокращении около 22% текущих вакансий. Ключ к решению проблемы — переобучение и повышение квалификации сотрудников. Российская программа «Цифровые профессии» — это государственная программа в рамках нацпроекта «Цифровая экономика», которая была запущена в 2019 году⁶. В рамках этой программы государство субсидирует обучение основным IT-профессиям и смежным специальностям.

Уже наблюдается другой тренд. После оформления нового специалиста на работу перед компаниями стоит задача

⁵ Искусственный интеллект позволит создать 170 млн новых рабочих мест // Русбейс URL: <https://rb.ru/news/ai-career-world/> (дата обращения: 20.05.2025).

⁶ Можно ли отучиться на айтишника за счет государства в 2025 году, проект «Цифровые профессии» // KEDU каталог образования URL: <https://kedu.ru/press-center/articles/otuchitsya-na-aytishnika-za-schet-gosudarstva/> (дата обращения: 20.05.2025).

поддерживать и развивать его навыки в условиях постоянных технологических изменений. То есть традиционные корпоративные курсы часто оказываются недостаточно эффективными и быстро теряют интерес сотрудников. На смену им приходят адаптивные обучающие платформы, где ИИ индивидуализирует контент под текущий уровень знаний и карьерные амбиции сотрудника.

Или персонализированные LMS-системы определяют оптимальную последовательность изучения модулей, рекомендуют темы для добавления или исключения, учитывая результаты тестов, интересы и опыт сотрудника. Это повышает вовлеченность, поскольку человек ощущает, что обучение действительно «под него», а не просто набор универсальных курсов.

Еще более инновационный формат — внутренние «маркетплейсы талантов». Сотрудник создает цифровой профиль с достижениями, интересами и карьерными планами. ИИ анализирует открытые вакансии и проектные задачи, подбирает подходящих кандидатов и одновременно советует, какие навыки стоит развивать для дальнейшего роста. Это способствует гибкой внутренней мобильности и снижает отток сотрудников, а не способствует их увольнению, поскольку они видят перспективы развития внутри компании.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Опасения оправданы для рутинных профессий (бухгалтеры, операторы колл-центров), где автоматизация достигла 40–60%. Однако ИИ создает новые специальности. Критически важна переподготовка кадров с учетом стремительного развития технологий.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Опасения не беспочвенны, поскольку ИИ уже берет на себя некоторые виды работ. Но пример предыдущих

технологических революций показывает, что технологии скорее не разрушают полностью, а трансформируют рынок труда. Каждая технологическая революция (паровая, электричество, компьютеры), помимо уничтожения старых рабочих мест, создавала новые, просто в иных сферах.

В краткосрочной перспективе существует риск замены сотрудников на ИИ-агентов там, где работа основана на повторяющихся рутинных действиях (например, операторы колл-центров, программисты уровня junior, сотрудники, занимающиеся первичной обработкой и вводом данных).

Однако в долгосрочной перспективе картина не такая однозначная. ИИ может привести к появлению целых новых отраслей и, соответственно, новых вакансий. По различным оценкам, ИИ технологии могут добавить 1–2% к ежегодному росту ВВП к 2040-м годам, а это подразумевает в том числе создание рабочих мест для реализации этого роста.

Таким образом, риск потери работы реален для отдельных профессий, но не является тотальной угрозой занятости человечества в целом. В итоге проблема для молодежи состоит не столько в массовой безработице, сколько в грамотном и заблаговременном приобретении навыков, дополняющих возможности ИИ.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Беспокойство молодежи относительно потери работы из-за развития искусственного интеллекта вполне обоснованно и имеет под собой реальные основания. История технологического прогресса показывает, что автоматизация всегда приводила к исчезновению определенных «устаревших» профессий. Однако важно понимать, что это процесс не новый — он происходил на протяжении всей истории человечества.

С одной стороны, искусственный интеллект действительно может заменить человека в рутинных, повторяющихся задачах, особенно в сферах обработки данных, производства и обслуживания. Уже сейчас мы видим, как автоматизация меняет рынок труда: роботы заменяют рабочих на производствах, чат-боты — операторов колл-центров, алгоритмы — аналитиков в некоторых областях.

С другой стороны, развитие ИИ создает новые возможности и профессии. Появляются специалисты по машинному обучению, разработчики нейросетей, этические консультанты по ИИ, специалисты по работе с большими данными. История показывает, что технологический прогресс всегда сопровождался не только исчезновением старых профессий, но и появлением новых, более сложных и требующих высокой квалификации.

Ключевым фактором адаптации к этим изменениям становится постоянное обучение и развитие новых навыков. Молодым людям необходимо ориентироваться на профессии, требующие творческого подхода, эмоционального интеллекта и способности к нестандартному мышлению. Именно эти качества пока остаются вне компетенции искусственного интеллекта.

Таким образом, хотя опасения молодежи оправданы, они не должны перерасти в панику. Вместо этого стоит рассматривать развитие ИИ как возможность для профессионального роста и переориентации на более сложные и интересные задачи, которые пока остаются в компетенции человека.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

На текущем уровне развития ИИ я такого риска не вижу. Всё абсолютно наоборот — не хватает персонала для развития экономики.

?

ВОПРОС

72% молодежи считают, что ИИ — новый «паровой двигатель», который даст развитие человечеству. Насколько разделяете этот образ?



Ольга Игоревна Пивень

Сооснователь и директор конференции ЦИПР

ИИ используется почти во всех областях нашей жизни и в этом смысле влияет не меньше, чем паровой двигатель или другие важные изобретения в истории человечества. Если говорить об экономической стороне дела, то нейросети берут на себя рутинные процессы, значительно повышая КПД и высвобождая человека для более творческих процессов. Такая оптимизация, безусловно, полезна.



Сергей Владимирович Володенков

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

То, что аналогию парового двигателя можно использовать, пожалуй, соглашусь. Насколько он принесет пользу в развитии человечества, здесь мог бы поспорить. С одной стороны, безусловно, становится удобнее во многих случаях, но быстрое получение информации отучает человека от самостоятельного ее поиска. Получение готовых решений отучает человека от самостоятельной выработки определенных представлений, способности делать самостоятельные выводы, потому что атрофируются эти способности, т. к. всегда есть тот, кто даст информацию, еще и сформулирует за меня определенную позицию. И в этом случае — да, легче и удобнее, в библиотеку можно не ходить, например. Здесь от самого человека зависит, будет искусственный интеллект способствовать развитию или упрощению, деградации человека, общества, цивилизации в целом.

Ответственность лежит на самом человеке, но, повторюсь, мое убеждение в том, что лень — двигатель прогресса. Я вижу больше потенциала для деградации, чем для прогресса, если мы говорим про современные модели применения искусственного

интеллекта для написания работы, поиска информации и так далее.

Основной риск — упрощение миропредставлений, потому что если тем же OpenAI ChatGPT пользуется 200–300 миллионов человек, это ценностно окрашенная сеть, которая дает готовые результаты, которые воспринимаются как объективные. Однако они необъективны. Большая языковая модель, на которой работает ChatGPT-5, обучена на ценностно окрашенных данных изначально. Как и любые человеческие данные также ценностно окрашены. Соответственно, эта окраска переносится на человека.

Таким образом, возможно глобальное управление мировосприятием, формирование единых универсальных представлений, если все начинают пользоваться одним инструментом. Здесь важно, во-первых, учитывать эту ценностную окрашенность, заниматься разъяснительной, просветительской работой, показывать, что ИИ не может быть объективным инструментом, плюс его галлюцинации, плюс определенные предубеждения, если будет угодно. Это все накладывает опечаток на формирование представлений человека.

Еще Фрэнк Уэбстер говорил, что мы живем не в мире, о котором у нас есть какая-то информация, напротив, мы обитаем в мире, созданном информацией. Если мы получаем информацию из псевдообъективного источника и выстраиваем свои модели представлений на ценностно окрашенных, упрощенных представлениях с элементами галлюцинаций и предубеждений, то мы можем прийти к тому, что просто неверно понимаем мир.

Второе, как я уже говорил, это упрощение жизни человека и мыслительных способностей. Мы слишком быстро можем получить готовый ответ, слишком просто получить ту информацию, которая нам необходима. Это не есть на самом деле хорошо. Конечно, с одной стороны, это хорошо, т. к. удобно и быстро, а с другой — не очень, потому что навыки могут потеряться и потом человек, оставшийся однажды, предположим, в результате глобального сбоя или кибератаки, без своего умного ассистента, окажется просто беспомощным.

Дальше, если мы говорим о столкновении цивилизаций по Хантингтону, в моем понимании, мы приходим к тому, что искусственный интеллект у каждой цивилизации будет иметь свои ценностно-смысловые характеристики, и люди будут пользоваться вот этими объяснительными моделями, которые

будут очень сильно различаться. И в оценке реальности, и исторических событий, и многих фактов действительности. В конечном итоге, это еще глубже может усугублять и углублять цивилизационные расколы.

К рискам можно отнести и слишком серьезное вовлечение искусственного интеллекта в ключевые сферы жизнедеятельности государства и общества в силу того, что любая техническая неисправность, глобальный сбой, вирус, выход искусственного интеллекта из подчинения, из-под контроля чреват тем, что человек остается голым, беспомощным. Он не способен продолжать свое эффективное функционирование в условиях, когда его ассистент отказывается ему помогать и снабжать его уже привычным образом необходимой информацией, необходимыми решениями, необходимыми предложениями, маршрутами, справками, расписаниями и так далее.

Можно говорить и об алгократии, алгоритмическом управлении. Если оно теневое и непонятное, я не исключаю того, что умные алгоритмы могут быть использованы в злонамеренных целях для злоупотребления возможностями использования технологий ИИ определенными технократическими элитами.

Если искусственный интеллект начинает играть серьезную роль в жизни человека, то у корпораций, у бенефициаров, которые владеют алгоритмами, возникает возможность цифровой депривации. Искусственный интеллект отказывается пропускать вас в лифт, в транспорт, в банк, на место работы, потому что бенефициар считает вас неблагонадежным. И, наконец, возникают алгоритмические паноптикумы.

Мы помним Иеремию Бентама, Мишеля Фуко, их концепции паноптикума. Сегодня в качестве невидимого наблюдателя может выступать алгоритм: он невидим, он наблюдает, и он же способен применять санкции в отношении тех людей, которые, по его параметрам, являются опасными, неприемлемыми и т. д. Формирование алгоритмических паноптикумов технологически возможно уже сегодня в масштабах стран или всего человечества, если рассматривать совсем уж такой глобальный сценарий.

Я не исключаю вероятность реализации такого сценария. Считаю это риском и даже угрозой социально-политического и гуманитарного характера. Мог бы рассказывать об этом часами, на самом деле, но вот, наверное, выделил самое ключевое.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Потенциал к этому у технологии, очевидно, есть. Хотя я бы назвал ее новым интернетом.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Сравнение ИИ с паровым двигателем — скорее выражение завышенных ожиданий. И даже оно выглядит скромно на фоне других аналогий, например, открытия ядерной энергии. Важно понимать, что технология ИИ в том или ином виде существует уже с 1960-х годов. Мы используем ее более полувека, пусть и на разных уровнях зрелости, в разных сферах.

Сегодня развитие ИИ зачастую движется по траектории «технология ради технологии». Часто сложно получить четкие ответы на вопросы: какую конкретную задачу решает та или иная система? Насколько она экономически или энергетически эффективна? Будь то большие языковые модели, машинное обучение или интеллектуальные агенты.

Тем не менее, существуют области применения, где ИИ действительно может дать серьезный рывок. Например, в разработке новых материалов, моделировании химических процессов, ускорении анализа больших объемов данных, автоматизации отдельных процессов. Однако и в этих случаях ИИ остается лишь дополнением к человеческому интеллекту, а не его заменой.

К сожалению, большая часть обсуждений вокруг ИИ сегодня носит популистский характер. Масштабные обещания, громкие лозунги и модные стартапы зачастую преувеличивают реальную пользу технологии и отвлекают внимание от ее реальных возможностей.

Таким образом, ИИ это не есть «паровой двигатель», но вполне может стать двигателем прогресса как вспомогательный инструмент. При этом его ценность будет зависеть от того, насколько осмысленно и ответственно мы научимся его применять.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Абсолютно согласен. ИИ действительно становится двигателем новой промышленной революции. Как паровой двигатель в свое время позволил механизировать производство, так ИИ сейчас меняет парадигму труда, коммуникации, принятия решений. Он открывает возможности для персонализированного образования, точной медицины, интеллектуальных городов.

Однако, как и в случае с паровым двигателем, успех зависит от того, насколько быстро и качественно общество адаптируется к этим изменениям. Россия имеет все шансы стать лидером в этой трансформации, особенно в тех отраслях, где уже есть компетенции, например в IT, образовании и госуправлении.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Уместное сравнение: паровая революция автоматизировала физический труд, ИИ трансформирует интеллектуальный. Однако риски масштабнее: неумелое использование влияет на приватность, психическое здоровье.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Искусственный интеллект уже демонстрирует трансформирующее воздействие на науку, экономику и повседневную жизнь, сопоставимое с ключевыми изобретениями прошлого. В целом, аналогия с паровым двигателем довольно близкая. ИИ — это универсальный инструмент, применимый почти повсеместно, имеющий трансдисциплинарный характер и в промышленности, и в науке. Например, в молекулярной биологии инструменты ИИ (тот же AlphaFold) уже решают задачу прогнозирования структуры белков, что ранее занимало десятилетия.

Однако аналогия справедлива еще и потому, что паровой двигатель, помимо индустриального скачка, принес и социальные потрясения (тяжелый труд, урбанизация, появление новых рисков). Эпоха ИИ также порождает новые вызовы — от упомянутых выше этических проблем до изменения рынка труда.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Сравнение искусственного интеллекта с паровым двигателем как революционной технологией, изменившей мир, действительно имеет под собой серьезные основания. Подобно тому, как паровая машина способствовала промышленной революции и кардинально изменила производственные процессы, так ИИ сегодня трансформирует практически все сферы человеческой деятельности — от медицины до искусства, от образования до производства.

Подобно паровому двигателю, который дал толчок развитию машиностроения и транспорта, искусственный интеллект открывает новые горизонты в автоматизации, анализе данных и создании интеллектуальных систем. Он становится универсальным инструментом, способным решать задачи, которые ранее были недоступны для массового применения из-за их сложности или трудоемкости.

Однако важно понимать, что, как и паровой двигатель в свое время, ИИ несет не только возможности, но и вызовы. Промышленная революция привела к социальным изменениям, перестройке экономики и появлению новых профессий (например, эколог). Точно так же развитие искусственного интеллекта потребует от общества адаптации, переобучения и переосмысления многих привычных процессов.

При этом аналогия с паровым двигателем в чем-то неполна. ИИ является более универсальным и гибким инструментом. Если паровая машина была в первую очередь источником механической силы, то искусственный интеллект способен решать задачи в самых разных областях, от творчества до научных исследований. Это делает его влияние на развитие человечества еще более масштабным и многогранным.

Таким образом, сравнение ИИ с паровым двигателем как символом технологической революции вполне оправданно.

Однако важно помнить, что это не просто аналог прошлого, а качественно новый инструмент развития, который может привести к еще более глубоким изменениям в обществе и технологиях.

Блок 2

Искусственный интеллект и страна

? ВОПРОС

Какие функции ИИ сегодня наиболее востребованы в России для бизнеса и органов государственной власти?

**Сергей Владимирович Володенков**

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Самое первое — это сбор, обработка, анализ, интерпретация информации, связанной с анализом существующих глобальных трендов, формированием глобальных трендов внутри государства, в том числе национальных внутри государства, на уровне государственных трендов. То есть это аналитико-прогностическая задача, когда мы можем проанализировать и дать прогнозы, сценарии развития ситуации.

Во-вторых, безусловно, искусственный интеллект позволяет при работе с big data адаптировать некоторые управленческие решения к тем взглядам, ожиданиям, представлениям людей, которые существуют у населения.

Получить это на основе анализа цифровых следов, в принципе, вполне возможно, и некоторые это делают, как на выборах Трампа 2016 года или при Брексите. Зная, чем дышит человек, на что он надеется, чего он ожидает, можно использовать искусственный интеллект для того, чтобы законы были более адаптированы к общественным, массовым, гражданским запросам. Мы можем понимать, чем живет сегодня общество, чем живет сегодня народ, чтобы оперативно отслеживать негативные тренды и вносить некие управленческие коррективы в процессы государственного управления. Вот самое основное.

**Андрей Алексеевич Митюков**

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

В бизнесе востребованы автоматизация клиентского сервиса (чат-боты), рекрутинг и оценка персонала, маркетинговая

аналитика, прогнозирование спроса, оптимизация логистики. В госсекторе — цифровые госуслуги, анализ обращений граждан, поддержка принятия управленческих решений, мониторинг социальных рисков.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Исходя из имеющегося опыта, можно сделать вывод, что в российском бизнес-секторе наибольшим спросом пользуются ИИ-решения, связанные с обработкой больших данных, автоматизацией документооборота и типовых бизнес-процессов.

Типичными задачами при этом являются прогнозирование спроса, разработка таргетированной рекламы, анализ паттернов в поведении клиентов, автоматическое распознавание и заполнение бухгалтерских и юридических документов и т. д.

Также быстро растет интерес к генеративным возможностям ИИ, включая генерацию текста, изображений, видео для маркетинга, дизайна и других задач. Большое количество компаний внедряют чат-ботов на базе ИИ, чтобы круглосуточно обслуживать клиентов в колл-центрах и службах поддержки.

В производственном секторе востребованы прогнозирующие модели для техобслуживания (предсказывать поломку оборудования) и компьютерное зрение для контроля качества продукции и соблюдения требований безопасности на производстве.

Для органов государственной власти приоритетны ИИ-функции, связанные с повышением эффективности услуг и общественной безопасности. В сфере госуправления широко внедряются интеллектуальные помощники и автоматизированные системы (яркий пример — сервис «Госуслуги»).

Решения на основе ИИ востребованы для оцифровки бумажных архивов, протоколов заседаний, судебных решений и т. д.

Государство также является одним из главных заказчиков на внедрение ИИ в сфере безопасности (системы видеонаблюдения с распознаванием лиц и номеров, аналитика прогнозирования преступности и т. д., см. вопрос 6 блока 2).

В области здравоохранения государство тестирует внедрение ИИ для помощи врачам: анализ медицинских изображений, распределение потока пациентов по степени срочности (например, для анализа КТ легких во времена COVID-19).

Таким образом, и бизнес, и власть ожидают от ИИ прежде всего эффективности и автоматизации. Бизнес — чтобы лучше знать клиента, сократить издержки и ускорить процессы (продвинутая аналитика, рекомендации, голосовые/визуальные сервисы). Государство — чтобы улучшить цифровые сервисы для населения и контроль инфраструктуры (цифровой документооборот, чат-боты для граждан, умная безопасность). Эти функции ИИ уже доказали свою полезность, поэтому являются самыми востребованными в России на сегодняшний день.



Иван Андреевич Никанов

Заместитель руководителя Исследовательского центра в сфере ИИ Университета Иннополис

Востребованы традиционные прогностические функции вроде анализа поведения клиентов, прогнозирования спроса и динамики цен. В последние годы значительно выросла популярность ИИ-агентов на базе больших языковых моделей для автоматизации универсальных задач, включая сортировку обращений, генерацию документации, а также обработку клиентских запросов в чатах и колл-центрах.

Кроме того, активно развиваются решения для автоматизации документооборота и проверки нормативно-правовых актов. Сохраняется интерес к системам компьютерного зрения, например для мониторинга объектов городской инфраструктуры, контроля за строительством и обеспечением безопасности. Растет спрос на интеллектуальные системы поддержки принятия решений, которые помогают анализировать большие массивы данных в городском управлении, здравоохранении и образовании.

?

ВОПРОС

С какими основными трудностями сталкиваются субъекты РФ при внедрении ИИ (например, инфраструктурные, кадровые, нормативные)?

Ключевые выводы:

1. Несовершенство инфраструктуры.

Описание: для эффективного ИИ нужны вычислительные мощности (центры обработки данных, специализированные ускорители) и большие объемы данных. В большинстве регионов нет своих мощных дата-центров уровня столичных, приходится либо арендовать облако, либо ограничиваться маломасштабными решениями.

2. Санкции и отсутствие производства чипов.

Описание: дефицит оборудования усугубился санкциями: Россия не производит свои чипы для машинного обучения, а импортировать новейшие GPU затруднительно.

3. Фрагментация данных между различными ведомствами и регионами.

Описание: фрагментация данных затрудняет их интеграцию и эффективное использование. Более того, некоторые усилия возможно объединить между ведомствами и регионами.

4. Дефицит кадров.

Описание: отмечается наличие дефицита кадров в данной сфере. Более того, бизнес не делится данными с вузами, что затрудняет подготовку специалистов.

5. Нормативные и бюрократические сложности

Описание: пока нет четких стандартов и регламентов на внедрение ИИ в госорганах, каждый регион идет своим путем, не всегда оптимальным. Руководство боится правовых рисков (непонятно, кто будет нести ответственность, если ИИ-система ошиблась). Кроме того, процедуры госзакупок зачастую не обладают гибкостью, необходимой для развертывания ИИ-

решений. Долгие согласования, низкая скорость принятия решений по новым технологиям также не идут на пользу делу.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

К основным трудностям внедрения ИИ в субъектах РФ можно отнести:

1. Несовершенство инфраструктуры. Для эффективного ИИ нужны вычислительные мощности (центры обработки данных, специализированные ускорители) и большие объемы данных. В большинстве регионов нет своих мощных дата-центров уровня столичных, приходится либо арендовать облако, либо ограничиваться маломасштабными решениями. Более того, дефицит оборудования усугубился санкциями: Россия не производит свои чипы для машинного обучения, а импортировать новейшие GPU затруднительно.
2. Недостаток квалифицированных кадров. Высококласные специалисты по ИИ сосредоточены в нескольких центрах (Москва, Петербург, Казань, Новосибирск). Регионы зачастую испытывают острый недостаток в специалистах необходимого профиля для развертывания и сопровождения ИИ-систем. Найм консультантов из крупных городов обходится слишком дорого. Кроме того, цифровая грамотность руководителей на местах не всегда достаточна, управленцы зачастую не понимают возможности ИИ и не поддерживают проекты.
3. Недостаток качественных данных. В регионах часто нет необходимых датасетов для машинного обучения. Данные не собираются системно, разрознены между ведомствами, плохо очищены. Частные компании не делятся данными друг с другом, не предоставляют данные вузам для подготовки специалистов. Подавляющая часть наиболее современных ИИ-моделей — зарубежные и обучались на англоязычных данных, что ведет к ограничениям и искажениям. Это приводит к необходимости зачастую начинать с нуля (сбор данных,

их очистка, разметка и т. д.). Это долгая и дорогая задача.

4. Недостаток финансирования. Внедрение ИИ решений на уровне региона требует значительных инвестиций и не всегда дает быстрый эффект. Особенно малый и средний бизнес чаще всего не может позволить себе эксперименты с ИИ и тем более разработку новых решений. Государственным учреждениям тоже проблематично выделить средства на что-то новое кроме приоритетных социальных статей. Более того, если проект требует закупки импортного ПО или оборудования, возникают сложности с оплатой из-за санкций.
5. Нормативные и бюрократические сложности. Пока нет четких стандартов и регламентов на внедрение ИИ в госорганах, каждый регион идет своим путем, не всегда оптимальным. Руководство боится правовых рисков (непонятно, кто будет нести ответственность, если ИИ-система ошиблась). Кроме того, процедуры госзакупок зачастую не обладают гибкостью, необходимой для развертывания ИИ-решений. Долгие согласования, низкая скорость принятия решений по новым технологиям также не идут на пользу делу.



Иван Андреевич Никанов

Заместитель руководителя Исследовательского центра в сфере ИИ Университета Иннополис

Субъекты РФ сталкиваются с рядом серьезных трудностей при внедрении ИИ. Одной из ключевых проблем является сильная фрагментация данных между различными ведомствами и регионами, что затрудняет их интеграцию и эффективное использование. Недостаток актуальных, чистых и машиночитаемых наборов данных препятствует успешному обучению и развертыванию ИИ-систем.

Существует нехватка мощных вычислительных ресурсов, таких как GPU-серверы и кластеры, что ограничивает возможности для выполнения сложных моделей и анализа больших данных. Высокая стоимость облачных решений и недостаточная инфраструктурная поддержка также становятся серьезным барьером для малых и средних предприятий и для региональных властей.

Не меньшую проблему представляет кадровый дефицит. Отток квалифицированных специалистов в столичные регионы и недостаточная подготовка кадров на местах тормозят развитие ИИ-инициатив в регионах. Вдобавок отсутствие единых нормативных стандартов и недостаточная правовая база создают неопределенность в вопросах защиты данных, этики и ответственности за использование ИИ, что также замедляет процесс внедрения



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Основные трудности, с которыми сталкиваются субъекты РФ при внедрении ИИ, включают инфраструктурные, кадровые и нормативные проблемы.

Инфраструктурные трудности связаны с недостаточным развитием информационных и коммуникационных технологий, а также с отсутствием или низким качеством связи и доступа к интернету. Это затрудняет использование ИИ в регионах и ограничивает возможности для внедрения новых технологий.

Кадровые проблемы связаны с нехваткой квалифицированных специалистов в области ИИ, особенно в регионах. Это затрудняет поиск и привлечение талантливых сотрудников, а также обучение и развитие существующих кадров.

Нормативные трудности связаны с отсутствием или недостаточной разработкой законодательства и нормативных актов, регулирующих использование ИИ в регионах. Это затрудняет создание благоприятных условий для развития ИИ, защиту прав и интересов субъектов, а также обеспечение безопасности данных и информации.

Для успешного внедрения ИИ в регионах необходимо преодолеть эти трудности и создать благоприятные условия для развития технологий и их использования в различных сферах жизни.

? ВОПРОС

Какие регионы России уже демонстрируют успешные практики внедрения ИИ? Что помогает этим регионам быть в авангарде?

**Алексей Степанович Паламарчук**

Генеральный директор NtechLab

Сегодня мы наблюдаем хорошую динамику внедрения ИИ в регионах России. Наша компания реализовала проекты в более, чем 70 регионах нашей страны в самых разных отраслях — от безопасности, ЖКХ и социальной сферы, до медицины и развития городской среды. Искусственный интеллект сегодня становится реальным и востребованным инструментом для решения целого спектра задач.

По соотношению вложенных ресурсов и полученной пользы ИИ-продукты, пожалуй, одни из самых эффективных. Зачастую решения не просто заменяют человека при выполнении типовых процессов, но и делают это быстрее, качественнее и исключая ошибки, продиктованные человеческим фактором.

Многие субъекты подходят к вопросу внедрения решений на основе ИИ комплексно, разрабатывая масштабные региональные программы. Здесь в качестве примера можно привести Москву, Санкт-Петербург, Нижегородскую и Новосибирскую области, Республику Татарстан, Краснодарский край.

Сегодня затраты на внедрение решений на основе ИИ окупаются во многом за счет предотвращения тех или иных происшествий и потерь. Кроме того, сегодня продукты на основе ИИ уже настолько распространены, что риск первооткрывателя фактически отсутствует: перед запуском решения всегда можно ознакомиться с практикой соседей-регионов и узнать, насколько они полезны, качественны и стабильны.

Поэтому работать ли с ИИ, внедрять ли его — это настолько очевидный выбор, как если бы мы сегодня спорили, стоит ли менять старый дисковый проводной телефон на современный смартфон.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Среди регионов России, демонстрирующих успешные практики внедрения искусственного интеллекта, можно выделить Москву, Санкт-Петербург, Татарстан, Свердловскую и Нижегородскую области. Эти субъекты активно участвуют в реализации федеральных проектов и программ в области цифровой трансформации и ИИ.

Москву отличает развитая цифровая инфраструктура, наличие высококвалифицированных кадров и плотное взаимодействие с ведущими научными центрами и ИТ-компаниями. Город также реализует пилотные проекты в рамках экспериментального правового режима, что способствует ускоренному внедрению ИИ-решений, особенно в сфере городского управления и здравоохранения.

Республика Татарстан благодаря технологическому хабу «Иннополис» и поддержке региональных властей демонстрирует активное развитие ИИ в сфере образования, транспорта и промышленности. Инфраструктурная обеспеченность, наличие профильных вузов и акцент на развитие цифровой экономики способствуют успеху региона.

Свердловская и Нижегородская области выделяются сотрудничеством между промышленными предприятиями и ИТ-сектором. Внедрение ИИ в производственные процессы и системы управления стало важным элементом повышения эффективности.

Общий фактор, объединяющий лидирующие регионы, — наличие стратегии цифрового развития, политическая воля на региональном уровне, поддержка научных и образовательных учреждений, а также участие в федеральных инициативах, включая создание экспериментальных правовых режимов для опробования ИИ-технологий в прикладных условиях. И, безусловно, рост объема инвестиций в технологию.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Среди российских регионов можно выделить несколько наиболее успешных примеров по внедрению ИИ. В первую

очередь это Москва, Санкт-Петербург и Московская область, где сосредоточены основные ресурсы и кадры, а власти сами иницируют масштабные проекты. Эти регионы создали благоприятную экосистему, включающую технопарки, акселераторы для AI-стартапов и крупных заказчиков (госорганизаций).

Также следует отметить Республику Татарстан (особенно Казань и Иннополис). В Казани традиционно сильна академическая научная школа, есть особая экономическая зона «Иннополис», предоставляющая льготы для бизнеса и широкие возможности для научных коллективов. Татарстан внедрил ряд сервисов с ИИ в госсекторе (электронное правительство с элементами ИИ, проекты умного транспорта).

За лидерами с некоторым отрывом следуют Ханты-Мансийский АО, Воронежская область, Республика Саха (Якутия), Ростовская область, Республика Башкортостан, Челябинская область и Алтайский край.

Перечисленные регионы являются экономически сильными, с развитыми отраслями промышленности, где ИИ приносит практическую пользу. Например, Ханты-Мансийский округ — центр нефтедобычи: там компании внедряют AI для сейсморазведки, оптимизации добычи (что поддерживается бюджетами предприятий).

Челябинская область и Башкортостан — индустриальные центры (металлургия, химия), где крупные корпорации (ММК, «Башнефть») инвестируют в промышленный ИИ.

Ростовская область — аграрно-промышленный регион, и там появляются AI-проекты в сельском хозяйстве (например, мониторинг посевов через спутники и дроны).

Воронежская область одной из первых внедрила алгоритмы распределения потоков пациентов по больницам).

Якутия (Саха) делает ставку на развитие стартапов («Якутская кремниевая долина» и другие акселераторы).

Ключевыми факторами успеха в развитии и внедрении ИИ технологий в указанных регионах, исходя из проведенного анализа, являются:

- Наличие научно-образовательных центров и кадрового резерва. Все упомянутые регионы имеют сильные университеты или филиалы, готовящие специалистов в области IT.

- Господдержка и прогрессивное руководство. В лидирующих регионах губернаторы и правительства непосредственно курируют цифровизацию.
- Экономическая база и партнерство с бизнесом. В богатых и промышленных регионах сами компании инициируют внедрение ИИ, оценивая перспективы и выгоду.

Таким образом, наиболее значимые результаты достигаются при наличии в регионе связки «вуз–корпорация–поддержка правительства».



Иван Андреевич Никанов

Заместитель руководителя Исследовательского центра в сфере ИИ Университета Иннополис

Республика Татарстан входит в десятку лидеров рейтинга регионов в области развития ИИ. Ключевыми факторами, способствующими успешному внедрению ИИ в этих регионах, являются наличие технических университетов и научных центров, развитая инфраструктура для исследований и разработок, а также поддержка со стороны региональных властей, готовых инвестировать в инновации и предоставлять налоговые льготы для стартапов в области ИИ. Эти регионы также активно привлекают федеральные и частные инвестиции, что позволяет значительно ускорить развитие технологий.

В Университете Иннополис есть такой центр — это один из 12 флагманских исследовательских центров в сфере ИИ в стране. Здесь разрабатываются и реализуются проекты, направленные на внедрение ИИ в различные отрасли — от здравоохранения и образования до промышленности и городской инфраструктуры. Причем не только в республике, но и в других российских регионах. Например, наш сервис AI Radiology CXR, который диагностирует на медицинских снимках патологии органов грудной клетки с точностью 93%, интегрирован в Единый радиологический информационный сервис и Единую медицинскую информационно-аналитическую систему Москвы и использовался в 450 тысячах исследований пациентов столичных клиник. А систему бесконтактного способа взвешивания свиней с помощью ИИ и компьютерного зрения используют в тульском агрохолдинге «Лазаревское».



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

В России есть регионы, которые успешно внедряют ИИ и демонстрируют хорошие результаты. Вот некоторые из них:

- Москва и Московская область активно используют ИИ для развития городской инфраструктуры, транспорта, образования и медицины. Здесь работают крупные IT-компании и научные центры, которые занимаются разработкой и внедрением новых технологий.
- В Республике Татарстан также активно внедряется ИИ в различные сферы жизни. Здесь создаются инновационные проекты, такие как умный город Альметьевск и цифровая долина Иннополис.
- Тюменская область использует ИИ для повышения эффективности нефтегазового сектора и развития туризма. Здесь работает центр искусственного интеллекта, который занимается разработкой и внедрением новых технологий.

Успешное внедрение ИИ в этих регионах обусловлено рядом факторов:

- наличие крупных IT-компаний и научных центров;
- поддержка со стороны государства и региональных властей;
- активное сотрудничество с бизнесом и образовательными учреждениями;
- развитие инфраструктуры и доступность интернета.



ВОПРОС

Насколько российские законы сегодня соответствуют темпам развития ИИ? Нужна ли специальная правовая база?

Ключевые выводы:**1. Проактивное, а не реактивное законодательство.**

Описание: ИИ развивается настолько быстро, что потребность во внесении поправок в законодательные акты будет лишь нарастать, поэтому в текущей ситуации целесообразно проводить не реактивную деятельность, а проактивную, создавая логику и рамки развития ИИ.

2. Отсутствие всеобъемлющего закона об ИИ.

Описание: большинство экспертов отмечают потребность в появлении закона, который в полной мере охватит развитие ИИ. Существующие акты и законы лишь частично охватывают деятельность ИИ, оставляя без внимания многие вопросы, в том числе ответственность за ошибки в его развитии. При этом важно, чтобы законы не проводили слишком жесткое регулирование, которое может затормозить инновации.

**Сергей Владимирович Володенков**

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Проблема любого законодательства, по моему субъективному мнению, заключается в том, что оно является не проактивным, а реактивным. В таком случае интенсивное развитие искусственного интеллекта будет требовать постоянного, ежедневного законотворчества, внесения изменений в законы, в правила, в процедуры регулирования.

Это придется делать постоянно и всегда с отставанием. Чем быстрее эволюция, тем больше отставание, потому что люди и институт законодательной власти имеют определенную инерцию в силу того, что это все-таки орган, институт, а не человек. И при этом эти люди имеют еще свою собственную инерцию объективно. Поэтому если мы остаемся в режиме

реактивного законотворчества, то мы будем все время отставать и все время менять, латать, трансформировать законодательство, что, естественно, не является оптимальным путем.

Если мы говорим про проактивное законотворчество, то, безусловно, здесь нужно обеспечить серьезный диалог законодателей со специалистами, практиками и учеными для того, чтобы заблаговременно формировать такое правовое поле, которое позволит успешно и эффективно развиваться умным технологиям, опять-таки в национальных интересах, и это поле будет задавать рамки и требовать минимальных корректировок. Я, может быть, идеалист в этом смысле, потому что я в Госдуме не сижу, законы не принимаю, но мне кажется, что переход к проактивному законотворчеству с продуманными рамками, которые позволят развиваться, не ориентируясь каждый день на то, чтобы прочитать новое постановление, наверное, более перспективно, скажем так. Но я прекрасно понимаю, что проактивное нормотворчество — это весьма серьезный вызов любой законотворческой системе, требующий серьезных усилий и компетенций.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Отсутствуют законы о регулировании генеративного ИИ и ответственности за ошибки.

Необходимо введение «регуляторных песочниц» для тестирования ИИ-решений.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

С учетом изложенного выше, по состоянию на сегодняшний день российское законодательство только частично охватывает сферу ИИ и уже заметно отстает от темпов его развития.

Прямого, всеобъемлющего закона об ИИ нет, есть только фрагментарное регулирование. В частности, имеется закон о персональных данных, закон «Об экспериментальных правовых

режимах в сфере цифровых инноваций в Российской Федерации» (позволяет временно опробовать ИИ-сервисы типа беспилотников в отдельных регионах), недавно вступил в силу закон о рекомендательных алгоритмах (обязывает крупные онлайн-сервисы давать пользователям возможность отключать персонализацию рекомендаций). Но все эти акты лишь точечно касаются отдельных аспектов. Например, вопрос ответственности за вред от действий ИИ-систем пока прямо не решен. По общим нормам отвечает либо изготовитель, либо оператор, но специальных правил нет.

Правоприменение также сталкивается с пробелами. Например, для ДТП с беспилотным автомобилем пока нет отдельного регулирования.

Таким образом, на текущий момент соответствие законов темпам развития ИИ условно удовлетворительное только в отдельных моментах. Прогресс ИИ ставит вопросы быстрее, чем появляются новые нормы. Нужна специальная нормативная база.



Иван Андреевич Никанов

Заместитель руководителя Исследовательского центра в сфере ИИ Университета Иннополис

Сейчас ученые, разработчики и исследователи в области ИИ обсуждают подходы к регулированию применения искусственного интеллекта в России. Обсуждаемая Концепция регулирования ИИ до 2030 года — важный шаг, но вопрос в том, насколько гибкой окажется правовая система. С одной стороны, слишком жесткое регулирование может затормозить инновации, с другой — отсутствие четких правил создает риски. Технологии развиваются стремительно, и правовая база всегда будет в роли догоняющего. Здесь важно соблюдать баланс между скоростью внедрения инноваций и предотвращением рисков от возникновения пробелов в законодательстве.

Пока Россия, как и многие страны, идет по пути адаптации существующих норм, Гражданского кодекса, законов о данных и авторском праве, но в перспективе потребуется более специализированная база. Ключевой вызов — найти баланс: не превратить регулирование в бюрократический барьер, но и не допустить правового вакуума в критических сферах вроде медицины, безопасности или госуправления.

? ВОПРОС

Какие международные инициативы или практики в сфере регулирования ИИ стоит учитывать России?

**Назир Заман Магси**

Преподаватель кафедры менеджмента Московского политехнического университета

В современных международных подходах к регулированию искусственного интеллекта наблюдается сдвиг от стратегии запретов к стратегии риск-ориентированного управления. Акцент смещается от ограничения технологии как таковой в сторону анализа ее применимости, уровня риска и потенциального воздействия на права и интересы человека. Именно этот подход представляется наиболее конструктивным и заслуживающим внимания при формировании российской системы регулирования ИИ.

Одним из наиболее развитых и системных нормативных документов в данной сфере является Акт Европейского союза об ИИ (EU AI Act)⁷, принятый в 2024 году. Он предлагает четырехуровневую классификацию ИИ-систем по степени риска: запрещенные (например, социальный рейтинг, манипуляции сознанием), высокий риск (системы, влияющие на здоровье, права или безопасность граждан), ограниченный риск (например, интерактивные интерфейсы), минимальный риск (например, рекомендательные алгоритмы для потребительских товаров).

Такой подход позволяет сочетать развитие технологий с защитой интересов человека и общества, не создавая ненужных барьеров там, где риск минимален. При этом согласно исследованиям⁸ данный акт содержит определение ИИ достаточно размыто и покрывает практически все алгоритмы.

⁷ The EU Artificial Intelligence Act Up-to-date developments and analyses of the EU AI Act // EU Artificial Intelligence Act URL: <https://artificialintelligenceact.eu/> (дата обращения: 21.05.2025).

⁸ Smuha N.A., Rengers E., Harkens A., Li W., MacLaren J., Piselli R., Yeung K. How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act // SSRN. – August 5, 2021. – 64 p. – URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991 (дата обращения: 21.05.2025).

В Великобритании реализуется децентрализованная модель регулирования, при которой отдельные отраслевые регуляторы получают полномочия формировать требования к ИИ в рамках своей сферы (здравоохранение, транспорт, финансы и пр.).⁹

Для сравнения, Китай использует более централизованную и жесткую модель. Алгоритмы подлежат обязательной регистрации, проходят контентную проверку, а поставщики обязаны обеспечивать контроль за содержанием и последствиями работы ИИ в реальном времени.

Таким образом, наиболее перспективным и сбалансированным направлением регулирования ИИ представляется интеграция риск-ориентированной модели с отраслевой спецификой. В рамках такого подхода ИИ-системы классифицируются по степени риска воздействия на человека и общество. Нормативные требования формулируются дифференцированно от строгих для критически важных сфер (медицина, правосудие), до рекомендаций для повседневных и потребительских технологий. Усилия должны концентрироваться не только на разработке и применении ИИ, но и на обеспечении прозрачности алгоритмов, обоснованности решений и этической ответственности.

Для России представляется актуальным заимствование логики приоритизации и прагматичности, а также включение в регуляторную рамку механизмов научной экспертизы, межотраслевого взаимодействия и открытого диалога между государством, бизнесом и научным сообществом.

В то же время важно осознавать, что попытка создать единый универсальный правовой акт, регулирующий искусственный интеллект во всех сферах, может привести к правовой неопределенности, коллизиям с уже действующими нормами и затруднению его практического применения. В этой связи целесообразнее сосредоточиться на поэтапном внедрении норм регулирования в рамках конкретных отраслей, где возможно четко определить специфику применения ИИ, уровень рисков и требования к надзору. Это обеспечит развитие регулирования ИИ не в виде отдельного закона, а как систему дополнений и адаптаций к существующему отраслевому законодательству, что обеспечит его гибкость, правовую точность и более быстрое внедрение в практику.

⁹ Artificial intelligence: ethics, governance and regulation // UK Parliament URL: <https://post.parliament.uk/artificial-intelligence-ethics-governance-and-regulation/> (дата обращения: 21.05.2025).



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

- Ethics Guidelines for Trustworthy AI (EC) — требования к прозрачности алгоритмов.
- NIST AI Risk Management Framework (США) — требования по кибербезопасности применения ИИ.
- Риски и проблемы при внедрении системы социального рейтинга Китая, критика за наблюдение за поведением, действиями человека.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

При формировании собственной системы регулирования ИИ России полезно обратить внимание на лучшие международные практики и инициативы в этой области:

Европейский опыт (AI Act¹⁰). Евросоюз сейчас на финальной стадии принятия всеобъемлющего регламента об ИИ — EU AI Act. Это первый в мире столь широкий законопроект, предлагающий классифицировать ИИ-системы по уровням риска и вводящий строгие правила для высокорискованных применений. В соответствии с ним в ЕС планируется прямо запретить ряд наиболее опасных практик ИИ: скрытное манипулирование людьми, внедрение социального рейтинга граждан, массовый биометрический надзор. Европейский подход также включает обязательство прозрачности — при общении с ИИ пользователю необходимо сообщить, что это не человек, а сгенерированный контент — он должен быть помечен.

Универсальные этические принципы (ЮНЕСКО, ОЭСР). UNESCO в 2021 году приняла «Рекомендацию по этическим аспектам ИИ¹¹», которую поддержала и Россия. Этот документ задает высокоуровневые ориентиры: уважение прав человека,

¹⁰ The EU Artificial Intelligence Act Up-to-date developments and analyses of the EU AI Act // EU Artificial Intelligence Act URL: <https://artificialintelligenceact.eu/> (дата обращения: 21.05.2025).

¹¹ Этические аспекты искусственного интеллекта // Unesco URL: <https://www.unesco.org/ru/artificial-intelligence/recommendation-ethics> (дата обращения: 21.05.2025).

принцип человеческого надзора, справедливость, недискриминация, ответственность разработчиков.

Организация экономического сотрудничества и развития (ОЭСР) тоже выработала в 2019 году принципы ИИ, которые поддержали многие страны. Россия, несмотря на то что не являлась членом ОЭСР, активно участвовала в обсуждениях. В документе делается акцент на надежности и безопасности ИИ на протяжении всего жизненного цикла, а также на международном сотрудничестве.

Китай в последние годы ввел ряд правил: закон о защите личных данных (аналог GDPR), регламент работы алгоритмических рекомендательных сервисов (требуется раскрывать основные принципы рекомендаций и обеспечивать отсутствие дискриминации), а с 2023 года — правила об управлении генеративным ИИ (обязательная цензура вредного контента, ответственность провайдеров).



Иван Андреевич Никанов

Заместитель руководителя Исследовательского центра в сфере ИИ Университета Иннополис

Из зарубежного опыта в сфере регулирования ИИ можно использовать подход UK AI Regulatory Framework, предложенный правительством Великобритании. Он основан на гибкости, отраслевой адаптации и стимулировании инноваций. В отличие от жестких нормативов других стран, британская модель делает ставку на стимулирование внедрения новых технологий, создание регуляторных «песочниц» и экспериментов в отдельных отраслях.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

России стоит учитывать следующие международные инициативы и практики в сфере регулирования ИИ:

- Инициатива ООН по искусственному интеллекту (IAI): направлена на разработку норм и принципов ответственного использования ИИ, а также на создание механизмов контроля и надзора за технологиями искусственного интеллекта.
- Принципы ОЭСР по ИИ: разработаны Организацией экономического сотрудничества и развития, касаются

прозрачности, безопасности и этичности использования ИИ. Включают рекомендации по регулированию данных, защите прав граждан и обеспечению безопасности ИИ.

- Глобальный договор ООН: призывает компании соблюдать принципы устойчивого развития, включая этические аспекты использования ИИ.
- Принципы «мягкого права»: «Принципы этики ИИ» IEEE и «Принципы ответственного развития ИИ» WEF предлагают рекомендации по разработке и внедрению ИИ, включая прозрачность, безопасность и этичность использования технологий.
- Международные стандарты и рекомендации: ISO/IEC 29100, IEEE 1824 и IEEE 1825, устанавливают требования к разработке, внедрению и использованию ИИ, включая безопасность, надежность и защиту данных.

Россия может использовать эти международные инициативы и практики для развития собственной стратегии регулирования ИИ, которая будет соответствовать международным стандартам и принципам ответственного использования технологий искусственного интеллекта.



ВОПРОС

Что должно измениться в системе взаимодействия науки, бизнеса и государства, чтобы ускорить внедрение ИИ в практику?

Ключевые выводы:**1. Исследования «на бумаге».**

Описание: зачастую многие решения остаются на бумаге из-за разрыва целей бизнеса, науки и государства. Необходимо приоритизировать общие цели.

2. Технологические трансфер-центры.

Описание: в России должны появиться трансфер-центры, которые возьмут на себя функцию коммуникации между наукой, бизнесом и государством, в том числе с целью применения различных разработок, которые могут являться готовым решением для интересантов.

3. Практико-ориентированные запросы государства.

Описание: со стороны государства следует совершать больше практико-ориентированных запросов бизнесу и науке, которые в том числе будут поддерживаться финансированием. Также есть потребность в поддержке малых инновационных предприятий и стартапов, которые зачастую предлагают прорывные решения.

**Андрей Алексеевич Митюков**

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Необходимо создать пространство для открытого диалога между всеми участниками.

Первое, что должно измениться — государство должно формулировать четкие, практико-ориентированные запросы и обеспечивать финансирование, ориентированное на конкретные результаты, а не на абстрактные исследования. Такой подход позволит направлять научные разработки в русло реальных бизнес-задач, повышая их применимость и

эффективность. Важным шагом станет развитие институтов технологического трансфера — механизмов, которые помогут быстро переводить научные инновации в коммерческие продукты и сервисы.

Второй аспект — активное вовлечение бизнеса в стандартизацию и тестирование ИИ-решений. Компании должны не только потреблять технологии, но и участвовать в формировании требований, пилотных проектах и обмене опытом.

Третья ключевая составляющая — поддержка малых инновационных предприятий и стартапов, которые часто являются источником прорывных идей. Государственные программы и венчурные фонды должны создавать благоприятные условия для их развития, обеспечивая доступ к инфраструктуре и рынкам.

Таким образом, для прорыва в применении ИИ России необходимо создать экосистему, где государство задает стратегию и финансирование, бизнес активно участвует в стандартизации и пилотах, а наука ориентируется на практическую применимость и технологический трансфер.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

- создание кросс-отраслевых консорциумов;
- введение обязательных модулей по машинному обучению в вузах;
- стимулирование бизнеса: налоговые вычеты в случае R&D в сфере ИИ или компенсация затрат на применение ИИ.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Для быстрого внедрения ИИ необходимо организовать более тесное и эффективное взаимодействие между научными учреждениями, бизнесом и государством. Сейчас это взаимодействие организовано недостаточно эффективно.

Трансфер технологий из науки в индустрию происходит крайне медленно по сравнению с КНР, например. Качественные

исследования зачастую остаются на бумаге из-за разрыва между целями науки и бизнеса.

Следует упростить вузам процедуру создания малых инновационных предприятий, предоставить налоговые льготы на первые годы для стартапов из университетов.

Хорошо зарекомендовали себя форматы консорциумов, где вуз и технологические компании совместно проводят исследования под прикладную задачу (например, квантовый консорциум).

Следует активнее стимулировать исследователей рублем за внедренные разработки. В настоящее время карьера ученого строится больше на публикациях, а не на внедрениях. Необходимо рассмотреть возможность изменения метрик и показателей эффективности.

Бизнесу также следует активнее вовлекаться в научную повестку, для чего необходимо расширять возможности для коммуникации, организовывать регулярные форумы, где исследовательские коллективы смогут презентовать свои ИИ-разработки потенциальным инвесторам и заказчикам.

Также нужны отраслевые лаборатории, когда корпорация открывает при вузе лабораторию по своему профилю (такая практика хорошо зарекомендовала себя в «Сириусе»). Это дает практическую задачу ученым и позволяет формировать кадры для бизнеса. Государство может софинансировать такие центры и лаборатории

Также требуются специалисты нового класса, которые одновременно смогут понимать задачи науки и бизнеса в области ИИ. Их подготовку могли бы взять на себя совместные образовательные программы вузов и компаний, например целевые магистратуры по ИИ

Целесообразно поощрять академическую мобильность. Ученому должно быть проще пойти работать в бизнес и вернуться, а представителю компании — прийти преподавать в вуз (такие практики развиваются, например, в ОЭЗ «Алабуга» в Татарстане, где инженерами-наставниками в «Алабуга Политех» работают ведущие специалисты предприятий-резидентов).



Иван Андреевич Никанов

Заместитель руководителя Исследовательского центра в сфере ИИ Университета Иннополис

Для успешного внедрения ИИ-решений в реальную практику требуются специальные интерфейсные структуры — своего рода технологические трансфер-центры, которые смогут эффективно соединять научные лаборатории с потенциальными заказчиками, переводя академические разработки в готовые рыночные продукты. Параллельно критически важно обеспечить исследовательским группам доступ к открытым и правильно анонимизированным государственным данным, которые являются бесценным ресурсом для обучения и тестирования ИИ-моделей.

При этом ключевая трансформация должна произойти в подходах всех участников процесса. Научному сообществу необходимо переориентироваться с публикационной активности на решение задач, актуальных для индустрии. Бизнесу следует отказаться от пассивного ожидания готового ИИ и активнее включаться в процесс разработки на ранних этапах. Государству же предстоит создать комплексную систему стимулов — от налоговых льгот до госзаказов, а также сформировать прозрачные механизмы валидации технологий, такие как единый реестр проверенных ИИ-решений с независимой экспертизой качества. Только такой комплексный подход позволит создать устойчивую экосистему внедрения искусственного интеллекта.

?

ВОПРОС

Какие достижения России в области ИИ считаете наиболее значимыми сегодня?



Ольга Игоревна Пивень

Сооснователь и директор конференции ЦИПР

На российском рынке появились две крупные модели для обработки естественного языка (LLM) — это YandexGPT и GigaChat. Что примечательно, обе были разработаны с меньшим объемом инвестиций, чем те, что потребовались для иностранного ChatGPT. Совсем недавно в Подмоскovie запустили сервис «Всесети», упрощающий процесс согласования документов при строительстве социальных объектов. В результате сроки рассмотрения сократились с года до двух недель. В дальнейшем предполагается масштабировать этот проект на федеральный уровень.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

У нас есть свои ИИ-команды, свои разработки, мы не сильно отстаем от США и Китая. И это огромное достижение.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Одним из ключевых и наиболее значимых и перспективных достижений России в области искусственного интеллекта является его применение в научных исследованиях, прежде всего в моделировании и разработке новых материалов и катализаторов.

Особое внимание заслуживает работа ученых Института органической химии им. Н. Д. Зелинского РАН (ИОХ РАН), где ИИ стал не просто инструментом анализа, а основой для научного прорыва. Среди наиболее значимых результатов — первая в мире нейросеть, определяющая молекулярную

формулу по фотографии вещества; автоматизированный спектрометрический анализ; технология 4D-катализа, позволяющая отслеживать химические процессы во времени и пространстве, и ряд других направлений.

Таким образом, именно интеграцию ИИ в фундаментальные научные задачи, такие как разработка новых материалов и катализаторов, можно считать одним из наиболее значимых и содержательных достижений России в этой области сегодня. Такие решения не только усиливают научный потенциал страны, но и создают базу для технологического рывка в химической промышленности и смежных отраслях.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Россия достигла заметных успехов в ключевых направлениях. Например, «Яндекс» и «Сбер» в 2025 году выпустили новые версии своих текстовых нейросетей — YandexGPT 5 Pro и GigaChat 2 Max, которые оптимизированы¹² для работы на более доступных устройствах без потери качества контента. Наша компания интегрировала генеративную сеть Yandex GPT¹³ в наш продукт для подбора персонала «Поток Рекрутмент», благодаря чему HR-специалисты и рекрутеры смогут формировать тексты вакансий, вопросы для собеседования, письма кандидатам через ИИ.

В России уже реализовано несколько заметных проектов по внедрению искусственного интеллекта в HRTech, которые показывают реальную эффективность и трансформируют процессы управления персоналом.

Например, это реализовали крупные компании — РЖД, «Ростелеком», «ДОМ.РФ». Они внедряют ИИ для оптимизации

¹² YandexGPT 5 Pro и GigaChat 2 Max против ChatGPT. Сравниваем три большие языковые модели // ROZETKED URL: <https://rozetked.me/articles/38394-yandexgpt-5-pro-i-gigachat-2-max-protiv-chatgpt-sravniyem-tri-bol-shie-yazykovye-modeli> (дата обращения: 20.05.2025).

¹³ HR-tech компания «Поток» внедрила генеративные нейросети в решение для поиска персонала // ИТ-маркетплейс CNews Market URL: https://www.cnews.ru/news/line/2024-03-13_hr-tech_kompaniya_potok_vnedrila (дата обращения: 20.05.2025).

HR-процессов, включая прогнозирование успешности кандидатов и персонализированное обучение сотрудников.

По данным исследований, 24% российских работодателей уже используют ИИ в HR, а 71% планируют внедрять новые инструменты в ближайшие годы. HRTech-компании и стартапы активно развивают технологии оценки soft skills через видеоинтервью, семантический анализ комментариев сотрудников и автоматизацию рутинных HR-задач.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Успехи России в области ИИ наиболее значимы в тех же областях, где исторически существовала сильная научная школа, прежде всего в компьютерном зрении и распознавании речи. Например, стоит отметить нейросеть от NTechLab, победившую в международных тестированиях NIST по точности распознавания лиц — в период 2018-2020 гг. я выполнял научное сопровождение проекта по тестированию стойкости алгоритмов NTechLab к спуфингу в интересах государственного заказчика.

Другое заметное достижение — системы обработки естественного языка для русского языка («Сбер» вместе с Институтом искусственного интеллекта AIRI и ЦРТ, «Яндекс» и др.) создали собственные крупные языковые модели, способные генерировать связный текст и отвечать на вопросы по-русски. Это заложило фундамент для развития отечественных чат-ботов, голосовых ассистентов, автоматизированного перевода и других NLP-приложений.

Также нельзя не отметить успехи в автономных системах: «Яндекс» развивает одну из передовых технологий беспилотных автомобилей, ее парк роботакси наездил свыше 10 млн километров в автономном режиме, в том числе в Иннополисе и в «Сириусе».

В сфере робототехники известен человекоподобный робот FEDOR (и его продолжения), а также военная платформа «Маркер». Они демонстрируют потенциал российских разработок автономных роботов.

Кроме того, значимы достижения в промышленном ИИ. Многие крупные предприятия («Сбер», «Газпром нефть», «Ростех») совместно с вузами внедрили системы предиктивной аналитики, цифровых двойников, оптимизации технологических процессов, которые заметно повышают эффективность производства.

Разработки Института системного программирования РАН по компьютерному зрению и доверенному ИИ используются в отечественных решениях для анализа видео на производствах. Согласно официальным данным, в 2022 году российские ученые зарегистрировали 905 результатов интеллектуальной деятельности в сфере ИИ (патенты, ПО и т. д.). Это показывает широту исследований.

Таким образом, сильные стороны РФ в ИИ сегодня — это алгоритмы мирового уровня в зрении и речи, успешное масштабирование этих алгоритмов (от лабораторий до реального применения), а также создание инфраструктуры (в РФ есть несколько мощных дата-центров для ИИ, например Christofari Neo от «Сбера», вошедший в топ-10 по вычислительной мощности в мире).



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Российские достижения в области искусственного интеллекта демонстрируют впечатляющий прогресс, особенно учитывая глобальную конкуренцию в этой стратегически важной сфере. Наиболее значимые разработки сосредоточены в нескольких ключевых направлениях, где российские ученые и компании смогли предложить оригинальные решения мирового уровня.

В области компьютерного зрения российские технологии занимают лидирующие позиции. Разработки компании VisionLabs признаны одними из самых точных в мире — их алгоритмы распознавания лиц используются не только в банковской сфере для биометрической идентификации, но и в городских системах безопасности. Особого внимания заслуживает платформа «Фейсграф», позволяющая анализировать эмоции и демографические характеристики с точностью до 98%. Эти технологии уже сегодня применяются в российской розничной торговле для изучения покупательского поведения и в транспортной инфраструктуре для обеспечения безопасности.

Российские достижения в обработке естественного языка также вызывают международный интерес. Языковые модели от «Яндекса», такие как YaLM 100B, демонстрируют результаты, сопоставимые с ведущими зарубежными аналогами. Особенностью российских разработок стала их адаптация к морфологической сложности русского языка и способность работать в условиях ограниченных вычислительных ресурсов. Система «Катибру» от МФТИ, например, показывает выдающиеся результаты в анализе научных текстов и генерации аннотаций, что открывает новые возможности для исследовательской работы.

В медицинской диагностике российские ИИ-системы добились прорывных результатов. Разработка Botkin.AI позволяет анализировать медицинские изображения с точностью, превышающей человеческие возможности. Система выявляет ранние признаки онкологических заболеваний, туберкулеза и COVID-19 по данным КТ и рентгенографии. В кардиологии алгоритмы, созданные в НИИ кардиологии Томска, прогнозируют риск инфаркта с точностью 93%, анализируя стандартные показатели ЭКГ. Эти технологии уже внедрены в десятках российских клиник и начинают экспортироваться за рубеж.

Особого упоминания заслуживают российские достижения в промышленном ИИ. Компания «Когнитивные технологии» разработала систему автономного управления сельскохозяйственной техникой, которая увеличивает урожайность на 15–20%. В нефтегазовой отрасли решения от «Газпром нефти» на основе ИИ оптимизируют добычу, экономя миллиарды рублей ежегодно. Российские алгоритмы предиктивной аналитики для промышленного оборудования, такие как «Предиктор» от «Росатома», значительно превосходят зарубежные аналоги по точности прогнозирования поломок.

Важно отметить, что многие российские разработки создаются с учетом специфики отечественной инфраструктуры и нормативной базы, что дает им преимущество на внутреннем рынке. При этом ряд технологий, особенно в области компьютерного зрения и медицинской диагностики, уже сегодня конкурируют на глобальном уровне. Дальнейшее развитие этих направлений может вывести Россию в число мировых лидеров искусственного интеллекта в специализированных нишах, что будет иметь стратегическое значение для технологического суверенитета страны.

Блок 3

Искусственный интеллект и будущее

**ВОПРОС**

Какие ключевые функции ИИ станут основными через 10–15 лет в безопасности (государственной и личной), промышленности, медицине, а также в освоении космоса? Что ожидать от ИИ в будущем?

Ключевые выводы:**1. Более широкое применение ИИ в области государственной и личной безопасности.**

Описание: в государственной и личной безопасности ИИ будет более широко представлен в системах мониторинга, предиктивного анализа и оперативного реагирования. Ключевые функции включают распознавание лиц и объектов, анализ поведенческих моделей, выявление аномалий, прогнозирование угроз и управление мерами безопасности в реальном времени. В сфере вооружений уже сейчас наблюдается стойкая тенденция к увеличению роли дронов и повышения автономности их действий. Следует ожидать, что интеллектуальные дроны возьмут на себя подавляющее большинство функций наблюдения и первичного реагирования на инциденты. ИИ научится не просто фиксировать угрозы, но и прогнозировать их с высокой точностью. Государственные системы безопасности будут использовать сложные нейросетевые модели для анализа миллионов параметров в реальном времени, выявляя потенциальные террористические акты или кибератаки еще на стадии подготовки.

2. Дальнейшая автоматизация.

Описание: в промышленности ИИ станет основой для автоматизации производственных и логистических процессов, цифрового управления технологическими линиями. Перспективы усиления участия технологии в области автоматического контроля качества продукции. ИИ станет основным инструментом для предиктивного анализа угроз, выявления кибератак и мониторинга социальной стабильности.

3. Право ИИ на принятие решения.

Описание: в ряде отраслей многие задачи будут полностью переданы технологиям ИИ, в том числе в части права на принятие решения.

4. ИИ как ассистент.

Описание: все большее признание от населения будет получать ИИ в качестве личного ассистента, которому возможно делегировать различные задачи.

5. Переход к суперинтеллекту.

Описание: переход к супер-ИИ — это не просто технологическая трансформация, а онтологический сдвиг, затрагивающий саму природу разума и бытия. Создается новый логос — нечеловеческий разум, который понимает закономерности быстрее, чем человек. Это система, которая понимает, учится, создает и решает задачи на уровне, недоступном человеку. Переход от «большого ИИ» к реальному супер-ИИ может произойти в пределах 5–7 лет.

6. Суперинтеллект и безопасность.

Описание: одной из ключевых опасностей является вопрос этики. Если суперинтеллект обладает искусственным разумом, возникает необходимость определить его права и ответственность. Например, может ли ИИ нести ответственность за свои действия? Эта проблема требует разработки новых правовых и моральных норм, которые пока отсутствуют. Другим важным аспектом является контроль за суперинтеллектом. Если ИИ начнет действовать вопреки интересам людей, как мы сможем его остановить? Отсутствие надежного механизма управления может привести к непредсказуемым последствиям и стать серьезной угрозой для глобальной безопасности.



Александр Юрьевич Назаров

Заместитель генерального директора Госкорпорации Ростех

Искусственный интеллект — не панацея, а мощный инструмент, который будет играть ключевую роль в оптимизации процессов, но не заменит человеческий фактор. Через 10–15 лет ИИ, скорее всего, станет стандартом в ряде отраслей, и в наших силах сделать его внедрение постепенным и контролируемым.

В промышленности он позволит сократить издержки за счет предиктивного обслуживания и автоматизации рутинных операций. В финансах — сможет улучшить оценку рисков и персонализацию услуг, но ключевые решения останутся за людьми. В кибербезопасности ИИ усилит защиту, но окончательные действия по-прежнему будут требовать экспертизы специалистов.

При этом с ростом возможностей растут и риски: автономное оружие, безработица, утечки данных и главное — потеря контроля. Кто ответит, если ИИ совершит ошибку? Разработчик, владелец или сам алгоритм? Безопасная интеграция ИИ возможна только при жестком регулировании, прозрачности алгоритмов и международной кооперации. Нам нужны не просто законы, а новая этика цифровой эпохи.

Уже сейчас ИИ меняет ИТ-индустрию, создавая новые профессии — промпт-инженер, ИИ-консультант. А завтра нейроинтерфейсы позволят управлять техникой силой мысли, а импланты — улучшать здоровье в реальном времени.

Главные риски здесь не в восстании машин, а в степени готовности общества и регуляторов. Безопасное внедрение ИИ требует консервативных сценариев и прагматичных шагов: четкие стандарты разработки и тестирования алгоритмов, постепенная интеграция с обязательным контролем человека, обучение сотрудников работе с ИИ, а не их замена.

Уверен, что ИИ не станет «властелином мира» — он останется инструментом, который можно использовать эффективно, без переоценки возможностей и недооценки рисков. Будущее не просто наступает — оно требует от нас готовности, адаптации к изменениям и внедрению прорывных технологий, облик которых мы же сами и определяем.

Мы можем научить ИИ ставить диагнозы точнее врачей и писать тексты грамотнее поэтов. Но мы не сможем запрограммировать в нем то, что делает нас людьми: способность задаваться

риторическими вопросами, любить вопреки логике, находить красоту в несовершенстве, наконец, быть этичными и человечными.

С развитием ИИ мы все яснее понимаем, что сила человека не в принятии безошибочных решений, а именно в том, что нельзя оцифровать, в том числе в нашей способности удивляться, любить и ошибаться, проявлять эмпатию, находить смысл за пределами логики.



Василий Юрьевич Бровко

Директор по особым поручениям Госкорпорации Ростех

Сейчас мы находимся на пороге события, сравнимого по значимости с появлением письменности, электричества или интернета. Речь идет о переходе от искусственного интеллекта к суперинтеллекту. Переход к супер-ИИ — это не просто технологическая трансформация, а онтологический сдвиг, затрагивающий саму природу разума и бытия. Илон Маск ставит своей целью через ИИ «понять истинную природу Вселенной» — это официальная миссия его компании. Философ Ник Бостром, автор книги *Superintelligence*, определяет суперинтеллект как «интеллект, который существенно превосходит лучшие человеческие умы практически во всех областях». В традиции Платона интеллект был путем к истине. Теперь мы создаем новый логос — нечеловеческий разум, который понимает закономерности быстрее, чем человек. Это не просто быстрый калькулятор или генератор текста. Это система, которая понимает, учится, создает и решает задачи на уровне, недоступном человеку.

Современные ИИ-системы умеют многое: распознают речь и изображения, пишут тексты, управляют машинами, анализируют огромные массивы данных. Но они всё еще ограничены. Это, по сути, узкоспециализированные алгоритмы, лишенные настоящего понимания, контекста или самосознания. Суперинтеллект — не просто «более умный ИИ». Это система, которая по своим когнитивным способностям превосходит человека во всех областях — от науки до искусства, от этики до интуиции. Это интеллект, который может создавать новые идеи, которые мы даже не в состоянии себе представить.

Переход к супер-ИИ будет проходить через три стадии:

Узкий ИИ — это наш сегодняшний день. Программы, решающие конкретные задачи лучше человека. Такие системы способны обрабатывать большие объемы данных и выполнять действия быстрее и точнее, чем человек, но они не могут выйти за пределы запрограммированных функций. Узкий ИИ не обладает самосознанием и не способен обобщать знания для решения проблем, которые не предусмотрены его алгоритмами, не способен выйти за пределы своей узкой специализации.

Системы слабого ИИ обучаются на основе данных, которые в них загружаются, и используют алгоритмы для поиска решений в определенной области. Например, они могут распознавать речь, переводить текст с одного языка на другой, делать рекомендации по контенту в интернете. Примеры — голосовые помощники (Siri, Google Assistant или «Алиса»), поисковые системы (Google, «Яндекс»), навигаторы (Google Maps или «Яндекс.Навигатор»).

Общий ИИ (Artificial General Intelligence, AGI) — система, обладающая универсальным интеллектом на уровне человека. Она может обучаться, обобщать и адаптироваться. Это система, способная выполнять широкий спектр задач.

В отличие от слабого ИИ, который ограничен узкими задачами, общий ИИ обладает универсальными возможностями. Они позволяют ему адаптироваться к новым условиям, учиться на собственном опыте и решать проблемы в разных областях. Он проектируется так, чтобы иметь способность понимать, учиться и принимать решения, используя обобщенные принципы, аналогичные человеческим. Это позволяет таким системам работать с разнообразными задачами, требующими логики и анализа на уровне, близком к человеческому, адаптироваться к новым ситуациям, а также развивать навыки без вмешательства человека.

Такой ИИ способен проводить анализ в любых областях — от математики и физики до искусства и социальных взаимодействий. Он может исследовать, давать советы, участвовать в принятии решений и даже разрабатывать новые подходы в науке и технике. Несмотря на потенциал, такой ИИ сейчас лишь в зачатке. GPT и другие нейросети выполняют различные задачи, но ограничены в своей способности к универсальному обучению и не обладают самосознанием.

Суперинтеллект (Artificial Superintelligence, ASI) — следующий скачок. Супер-ИИ — это гипотетическая система, обладающая интеллектом, который значительно превышает человеческие

способности во всех областях. В отличие от общего ИИ, супер-ИИ будет не просто выполнять те же задачи, а также генерировать новые подходы и решения, которые выходят за пределы возможностей человеческого разума. Такая система способна быстро адаптироваться, обучаться и совершенствовать алгоритмы, что позволяет ей решать задачи, которые кажутся слишком сложными для людей.

Супер-ИИ может изучать неизвестные до сих пор теории, разрабатывать новые технологии или предсказывать события с точностью, недостижимой для человека. Супер-ИИ решает задачи, которые находятся за пределами возможностей человека, и способен создавать новые концепции и идеи, непостижимые уму. Супер-ИИ способен к постоянному самосовершенствованию. Он может самостоятельно улучшать алгоритмы и учиться на собственных ошибках, таким образом эволюционируя и становясь мощнее с каждым шагом. В отличие от людей, он не ограничен возрастом, усталостью или биологическими факторами.

Сейчас мы стоим где-то между этапами. Еще в 2016 г. ИИ от DeepMind AlphaGo впервые победил чемпиона мира по игре го. Ранее это было невозможно с точки зрения алгоритмов. Точное количество разрешенных комбинаций в го оказалось больше числа атомов во Вселенной.

В 2020 году GPT-3 от OpenAI имела млрд параметров/соединений. GPT-4 в несколько раз мощнее. Для сравнения: человеческий мозг имеет более чем 100 трлн нейронов. В 2022 году ИИ от DeepMind AlphaFold предсказал структуру более 200 млн белков, что кардинально ускорило биомедицинские исследования. Это задача, которую ученые решали десятилетиями вручную.

ИИ развивается невероятно быстро. По прогнозам Meta AI, Google DeepMind и OpenAI, переход от «большого ИИ» к реальному супер-ИИ может произойти в пределах 5–7 лет. Внутри некоторых лабораторий уже тестируются системы, которые умеют рассуждать о целях без прямой постановки задачи. Это ключевой признак интеллектуальности — способность действовать, когда нет точных инструкций. Например, если ИИ в городской камере замечает необычное поведение толпы, он может не просто зафиксировать событие, а принять решение: уведомить оператора, изменить маршрут дронов, активировать экстренную подсветку. Это не просто распознавание — это инициатива.

Несмотря на бурное развитие, видимый прогресс нейронных сетей не может быть бесконечным. Уже сейчас технологии ИИ приближаются к своим пределам. Переход от узкого ИИ к общему, и, как следствие, к супер-ИИ потребует аналогичных по значимости с их изобретением прорывов в том, как мы создаем подобные системы.

Текущие принципы создания и обучения ИИ всецело полагаются на гигантские массивы данных, которые в большей своей части исчерпаны. Для создания AGI с применением текущих методов потребуется на порядок большее количество данных, создание которых не представляется возможным. В этой части развивается альтернативный путь обучения ИИ — сенсорное понимание мира. Как пример: Илон Маск разрабатывает технологию использования человеческого мозга как сенсора внешнего мира для обучаемого ИИ.

Также текущий тренд на увеличение мощности ИИ многократно увеличивает энергетические требования к нему. Так, разработки Amazon используют в том числе дата-центр, базирующийся на атомной электростанции мощностью в 960 МВт. Аналогично Microsoft ведет переговоры для перезапуска ядерной электростанции, чтобы обеспечить энергией свои разработки ИИ. Потенциальный AGI, основанный на текущих технологиях, потребует десятки гигаватт энергии. Для сравнения, человеческий мозг, при всей его вычислительной способности, потребляет не более 1 000 калорий в день. Преодоление данного разрыва в энергоэффективности потребует развития принципиально новых подходов к вычислениям — квантовых, органических, гибридных и прочих гипотетических.

Когда мы говорим о применимости ИИ в реальном мире, видеоаналитика — одна из самых ярких и наглядных сфер. Это уже не только камера наблюдения, а «глаз и мозг» цифровой эпохи. Видеоаналитика — не просто отрасль, где применим ИИ. Это полигон, на котором рождается реальный переход от ИИ к суперинтеллекту. Сегодня камеры видят, завтра они будут понимать и предсказывать, а послезавтра — принимать решения. Когда-то камеры просто смотрели. Сегодня они начинают думать.

Что уже умеет ИИ в видеоаналитике? Распознать лицо из 10 000 в толпе за доли секунды, определить подозрительное поведение в торговом центре или аэропорту, предсказывать аварии за секунды до столкновения на перекрестке. Но всё это

— лишь проявления узкого искусственного интеллекта. Он фиксирует факты, но не понимает их.

Современные системы искусственного интеллекта в видеонаблюдении уже впечатляют своими возможностями, но на фоне грядущего суперинтеллекта они выглядят ограниченными и уязвимыми. Их главный недостаток — отсутствие контекста. Алгоритм может распознать падение человека, но не способен понять, что за этим стоит: неудачный трюк на скейтборде, игра в догонялки или акт насилия. Неспособность к интерпретации — одна из фундаментальных проблем узкого ИИ.

Кроме того, системы все еще чувствительны к условиям съемки. Плохое освещение, нестандартный ракурс, дождь, дым, толпа — всё это, хоть и с низкой вероятностью, но способно сбить их с толку. Любое сильное отклонение от «знакомой» картины может привести к ошибке, потому что ИИ не умеет переносить опыт с одной ситуации на другую. Он заучивает шаблоны, но не понимает их суть.

Слабым звеном остается и неспособность к обобщению. Если ИИ обучен на кадрах из супермаркета, он бесполезен в метро. Ему придется «переучиваться» с нуля, тогда как человек способен интуитивно адаптироваться к новым условиям — именно это и станет ключевой задачей будущего суперинтеллекта.

Суперинтеллект меняет саму природу восприятия. Он не просто анализирует изображение — он понимает смысл происходящего. Это качественный скачок от «зрения» к «пониманию». Через некоторое время системы видеонаблюдения будут:

Понимать контекст. Они научатся интерпретировать действия, различать мотивы. Например, отличать драку от игры, стресс от агрессии, помощь от нападения. Предсказывать события. Суперинтеллект будет анализировать поведение в динамике и предугадывать потенциальные угрозы до их реализации. Это уже не просто безопасность — это предиктивная аналитика, или, как ее называют, pre-crime AI. Самообучаться на месте. Без дообучения и участия программистов такие системы смогут адаптироваться к новой среде: другой архитектуре, погоде, культурным особенностям. Работать мультиканально. Суперинтеллект объединит видео, аудио, текстовые данные, сигналы с IoT-датчиков в единую систему восприятия. Это даст

полную картину происходящего — как у человека, но в режиме реального времени.

То есть, мы перейдем от технической фиксации к семантическому зрению — пониманию смысла сцен. Система будет видеть не просто «человека в пальто», а «нервного преступника в стрессовой ситуации». Она сможет реагировать мгновенно, без задержек, еще до того, как человек осознает, что происходит. Более того, она сама будет ставить перед собой цели анализа. Если сегодня параметры задаются вручную, то завтра система сможет понять: здесь нужно усилить наблюдение, там — исключить ложные срабатывания.

И мы близки к этому. Уже сегодня искусственный интеллект NtechLab предотвращает преступления и находит пропавших людей. С 2017 года с помощью ИИ-решений NtechLab раскрыто более 20 тысяч преступлений в России. ИИ помогает врачам в диагностике инсультов. Искусственный интеллект NtechLab обработал 100 000 КТ-исследований головного мозга, оценка ИИ каждого из снимков заняла меньше минуты. ИИ от NtechLab не только распознает жизненные ситуации (например, курение, езду на самокате в неположенном месте), но и самостоятельно реагирует на них, например вежливо просит посетителей парков сменить локацию. По статистике, число нарушений в парках, где работает ИИ NtechLab, значительно сокращается.

Однако широкое распространение ИИ, даже узкоспециализированных, вносит значительные изменения в традиционный уклад жизни всего населения земного шара. Где ИИ используется с пользой, там обязательно появляется возможность для менее благих целей. Уже сейчас ИИ используется для мошенничества, обмана и прочих способов нарушения закона. Помимо преступлений, использование ИИ даже в относительно благих намерениях вызывает неоднозначные, зачастую поляризованные отклики у населения. Использование ИИ спецслужбами для слежки за людьми и потенциального предотвращения преступлений уже сталкивается с обвинениями в расовом профилировании и вторжении в личную жизнь. Появление суперинтеллекта, обладающего способностями к обучению и принятию решений, превосходящими человеческие, усугубит подобные проблемы, потенциально выступая как мощным драйвером прогресса, так и угрозой для существования человечества.

Одной из ключевых опасностей является вопрос этики. Если суперинтеллект обладает искусственным разумом, возникает необходимость определить его права и ответственность.

Например, может ли ИИ нести ответственность за свои действия? Эта проблема требует разработки новых правовых и моральных норм, которые пока отсутствуют.

Другим важным аспектом является контроль за суперинтеллектом. Если ИИ начнет действовать вопреки интересам людей, как мы сможем его остановить? Отсутствие надежного механизма управления может привести к непредсказуемым последствиям и стать серьезной угрозой для глобальной безопасности. Суперинтеллект может развиваться в направлениях, которые не будут учитывать интересы человечества. Например, супер-ИИ может игнорировать экологические последствия или права человека ради достижения своих целей. Само появление ASI сопряжено с вопросом — а может ли он существовать, совмещая принципы морали и человечности и свои сверхчеловеческие способности? Или в итоге мы получим не более чем искусственного человека, со всеми его недостатками и слабостями?



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Учитывая текущие темпы и направление технологического прогресса, можно говорить не о резкой революции в применении ИИ, а о постепенном усилении существующих функций. Прорывной рост возможен только при появлении новых вычислительных моделей или архитектур, что пока находится на стадии фундаментальных исследований. Поэтому горизонты 10–15 лет следует рассматривать как условные ориентиры, подлежащие регулярной переоценке. Более реалистичными для практического планирования являются 5-летние циклы.

Тем не менее с сохранением текущей динамики можно выделить ряд ключевых функций искусственного интеллекта, которые, вероятно, станут основными в различных отраслях.

В государственной и личной безопасности ИИ будет более широко представлен в системах мониторинга, предиктивного анализа и оперативного реагирования. Ключевые функции включают распознавание лиц и объектов, анализ поведенческих моделей, выявление аномалий,

прогнозирование угроз и управление мерами безопасности в реальном времени. В перспективе возможно создание интегрированных аналитических платформ, способных существенно повысить эффективность систем национальной и региональной безопасности.

В промышленности ИИ станет основой для автоматизации производственных и логистических процессов, цифрового управления технологическими линиями. Хорошие перспективы усиления участия технологии в области автоматического контроля качества продукции, оптимизации энергопотребления и управления логистическими цепочками с целью повышения эффективности и снижения издержек.

В области медицины основными функциями ИИ можно выделить раннюю диагностику заболеваний, персонализированный подбор терапии, моделирование биологических процессов. Особенно перспективным направлением является применение ИИ в разработке лекарственных препаратов.

В космической отрасли применение технологии будет развиваться в области автономной навигации и управления космическими аппаратами, а также анализа данных, полученных с телескопов, спутников и марсоходов.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Давайте будем честны: скорость текущей технологической революции такова, что мы абсолютно не понимаем, к чему она приведет через 10–15 лет. Человек так устроен, что строит линейные прогнозы на будущее, а мы имеем дело с самой крутой технологической экспонентой за всю историю. Поэтому мое видение будущего, скорее всего, тоже линейно.

Через 10–15 лет искусственный интеллект станет стратегическим партнером в управлении персоналом, выйдя за рамки автоматизации и став ключевым элементом экосистемы талантов. Он будет не просто обрабатывать данные, а участвовать в принятии решений, основанных на понимании человеческого потенциала и потребностей бизнеса.

Одной из центральных задач ИИ станет прогнозирование успешности кандидатов и удержания сотрудников. Это позволит HR переходить к превентивным действиям, повышая эффективность управления талантами. Развитие персонала станет полностью персонализированным: ИИ будет формировать индивидуальные траектории обучения и карьерного роста, адаптируясь под меняющиеся рынки и технологии.

В рекрутменте ИИ будет анализировать как профессиональные навыки, так и поведенческие характеристики. При этом такие системы будут сопровождаться строгими этическими стандартами, направленными на исключение предвзятости. Прозрачность, объяснимость решений и возможность их оспорить станут обязательными.

Еще одной важной функцией станет непрерывная оценка компетенций и вовлеченности внутри компании. ИИ поможет отслеживать развитие команд, предлагать меры по усилению культуры и оперативно закрывать дефицит знаний.

Каждый сотрудник получит личного цифрового HR-ассистента — ИИ, который будет сопровождать его на всех этапах жизненного цикла в компании.

Эти изменения потребуют от компаний технологической зрелости, соблюдения этических принципов и культуры доверия к технологиям. Те, кто закладывает основу для ответственного использования ИИ уже сегодня, станут лидерами в управлении талантами будущего.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

В сфере безопасности ИИ станет основным инструментом для предиктивного анализа угроз, выявления кибератак и мониторинга социальной стабильности. В промышленности алгоритмы ИИ найдут применение в оптимизации логистики, управлении энергопотреблением, в автономных роботизированных системах. В медицине акцент сместится на разработку персональных терапевтических стратегий и нейроинтерфейсов. В космосе — на создание автономных систем для анализа внеземных ресурсов.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

С большой долей вероятности через 10–15 лет ИИ станет центральным элементом систем как государственной, так и личной безопасности.

В сфере государственной безопасности интеллектуальные системы мониторинга будут в реальном времени анализировать видеопотоки, интернет-контент, поведенческие данные для предупреждения преступлений и угроз (в частности, на решение той же задачи направлен грант РНФ № 25-11-20034 «Технология раннего выявления угроз информационно-психологической безопасности учащихся образовательных организаций Федеральной территории "Сириус" на основе моделей и методов искусственного интеллекта»).

Уже сейчас закладываются основы умных городов, а к 2035 году такие системы станут еще более автономными и проактивными. Например, ИИ будет выявлять подозрительные действия на улицах или попытки кибератак и мгновенно оповещать соответствующие службы. В сфере вооружений уже сейчас наблюдается стойкая тенденция к увеличению роли дронов и повышению автономности их действий. Следует ожидать, что интеллектуальные дроны возьмут на себя подавляющее большинство функций наблюдения и первичного реагирования на инциденты.

В области личной безопасности широкое распространение получат домашние ИИ-ассистенты, следящие за жильем (распознающие пожар, протечку, проникновение посторонних), носимые устройства с ИИ, которые будут вести мониторинг окружения человека и сигнализировать о возможной опасности.

В промышленности ключевой функцией ИИ станет полная автономизация и оптимизация производства. Стоит ожидать появления умных производств, где ИИ будет управлять большинством процессов — от закупки сырья и управления цепочками поставок до контроля качества продукции. При этом предиктивная аналитика на базе ИИ будет стандартом — машины сами будут предсказывать, когда им потребуется обслуживание или настройка. ИИ также с высокой долей вероятности возьмет на себя дизайн и проектирование — генеративные модели смогут проектировать оптимальные конструкции деталей, схемы заводов и т.п.

Таким образом, главными функциями ИИ в индустрии будущего станут автоматизация и самооптимизация.

В медицине следует ожидать существенного усиления роли ИИ в сфере диагностики заболеваний. Наиболее вероятно, что она станет основной формой диагностирования, особенно для бедных слоев населения. Также значительно повысится эффективность персонализированного лечения. ИИ сможет агрегировать информацию о пациенте (медицинская история, генетика, образ жизни) и предлагать индивидуальные схемы терапии, подстраивая курс лечения под конкретного человека.

Также за счет ИИ должно существенно сократиться время, необходимое для разработки лекарств. В итоге медицина станет более точной, проактивной (предупреждать болезнь, а не только лечить) и доступной для широких слоев населения, в том числе в беднейших странах.

В освоении космоса ИИ, вероятно, возьмет на себя роль автопилота для большинства проектов. ИИ будет управлять полетом межпланетных станций, самостоятельно корректировать курс, исходя из обстановки, выбирать оптимальные траектории. При колонизации Луны или Марса роботы с ИИ будут заниматься строительством баз, добычей полезных ископаемых, обслуживанием инфраструктуры без прямой команды с Земли (поскольку связь с задержкой).

Спутниковые группировки будут использовать «роевой» интеллект, избегая столкновений, перестраиваясь под текущие задачи в режиме реального времени. В научных проектах анализ космических данных с помощью ИИ позволит находить новые объекты и аномалии быстрее, чем это происходит сейчас.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

В долгосрочной перспективе через 10–15 лет искусственный интеллект кардинально преобразует ключевые сферы человеческой деятельности, предложив принципиально новые подходы к решению сложных задач. В области безопасности произойдет переход от реактивных к предиктивным системам — ИИ научится не просто фиксировать угрозы, но и прогнозировать их с высокой точностью. Государственные системы безопасности будут использовать сложные нейросетевые модели для анализа миллионов параметров в реальном времени, выявляя потенциальные террористические акты или кибератаки еще на стадии

подготовки. На персональном уровне может получить широкое распространение умный антивирус, непрерывно анализирующий цифровое окружение человека и блокирующий фишинговые атаки до того, как пользователь успеет совершить опасное действие. Особое развитие могут получить технологии децентрализованной идентификации, где ИИ будет управлять персональными цифровыми паспортами, исключая возможность подделки и несанкционированного доступа.

В промышленности ИИ трансформирует всю цепочку, от проектирования до логистики. Появятся полностью автономные «цифровые двойники» заводов, способные самостоятельно оптимизировать производственные процессы в реальном времени с учетом тысяч переменных, от рыночного спроса до состояния оборудования. Квантово-гибридные алгоритмы позволят моделировать новые материалы с заданными свойствами без дорогостоящих физических экспериментов. В энергетике ИИ-системы обеспечат балансировку сложных сетей с преобладанием возобновляемых источников, прогнозируя генерацию и потребление с точностью до 99%. Особое значение приобретут промышленные киберфизические системы, где ИИ будет управлять симбиозом роботов, 3D-принтеров и человеческих операторов на опасных производствах.

Медицина будущего будет основываться на концепции непрерывного здоровья, где ИИ-ассистенты, анализируя данные носимых датчиков и геномную информацию, смогут выявлять заболевания на доклинической стадии. Персонализированные медицинские ИИ разработают индивидуальные схемы лечения с учетом молекулярного профиля пациента, а хирургические роботы нового поколения смогут выполнять операции с наноточностью под контролем нейросетей. Особый прорыв ожидается в нейропротезировании — интерфейсы мозг-компьютер с ИИ смогут восстанавливать утраченные функции движения и даже элементы когнитивной деятельности. В фармацевтике генеративные модели ускорят разработку новых препаратов с нескольких лет до месяцев, моделируя их воздействие на виртуальных пациентах.

Кроме того, в еще более долгосрочной перспективе ИИ может вывести на качественно новый уровень освоение космоса. Например, космические аппараты следующих поколений с внедренными автономными ИИ-системами смогут самостоятельно принимать решения в условиях задержки

сигнала, анализировать геологические данные на других планетах и выбирать оптимальные места для баз. В рамках лунных и марсианских программ ИИ возьмет на себя управление системами жизнеобеспечения, адаптируя их работу к изменяющимся условиям. Особое значение приобретут роботы-астронавты с ИИ, способные выполнять сложный ремонт в открытом космосе или вести разведку в опасных для человека условиях.

Эти преобразования потребуют пересмотра существующих этических норм и правовых рамок, но именно они определяют технологическое лидерство стран в середине XXI века. Россия имеет все возможности занять достойное лидирующее место в этом новом технологическом укладе.



Александр Александрович Любченко

Заведующий лабораторией искусственного интеллекта, ООО «ОЦРВ», филиал «Сириус», холдинг «РЖД-Технологии»

Хотелось бы начать с того, что в мировой практике продолжатся исследования и разработка технологий, приближающих человечество к достижению цели по построению общего, сильного искусственного интеллекта (AGI). Сегодня существуют разные прогнозы относительно временного горизонта такого события, и всё зависит от того, что же на самом деле представляет из себя AGI. За предстоящие 10–15 лет мы больше разберемся в этом вопросе. Теоретические предположения будут подтверждены на практике, будут сформулированы более современные критерии отнесения технологии к сильному ИИ. При этом не секрет, что внедряемые сегодня в различных отраслях цифровые сервисы на основе технологий ИИ — это отдельные реализации и решения задач, относящихся к слабому или узкому ИИ. Эти задачи объединены в технологические группы, например компьютерное зрение, распознавание и синтез речи, обработка естественного языка. Возможно, была бы не лишней отдельная группа, объединяющая задачи генеративного ИИ ввиду широкого распространения этих технологий сегодня.

В целом для ИИ-систем отраслевого назначения в грядущем десятилетии прогнозируется тенденция повышения уровня автоматизации интеллектуальных функций, выполняемых человеком. При этом хотелось бы упомянуть одно из ожиданий заказчика, которое часто декларируется сегодня, и, возможно, станет одной из ключевых функций, — это возможность

распределенного самообучения моделей под задачу. Такая функциональность предполагает автономность ИИ-сервисов в вопросах поддержания стабильного качества работы без необходимого вмешательства со стороны разработчика, учитывая отраслевую специфику. Также важной будет функция интерпретации решений с объяснением, на основе каких данных, фактов, суждений ИИ-модель прогнозирует, генерирует ответ и предлагает решение.

Переходя к вопросу личного использования ИИ-инструментов, есть любопытное наблюдение по генеративному ИИ. Как будто этот класс технологий быстрее распространяется и приходит в обыденную жизнь граждан, нежели промышленные реализации таких сервисов с внедрением в отраслях. В принципе, сегодня многие граждане могут без относительного труда воспользоваться доступными в сети Интернет ИИ-сервисами на платной и бесплатной основе. Не нужно задумываться о вычислительной инфраструктуре, интеграциях. Всё уже сделано, быстрая регистрация — и можно пользоваться умным чат-ботом или создавать изображения и даже видео. Это повышает доступность ИИ-сервисов для общества, постепенно формируется доверие к технологиям. Учитывая это, вероятно, уровень автоматизации интеллектуальных функций в обществе будет расти быстрее, нежели в отдельных отраслях, ввиду значительно большего количества пользователей и повышения доступности таких сервисов. Не исключено, что в таком случае востребованной функцией будет способность выявлять сгенерированный синтетический контент и прочие ИИ-ответы и, возможно, такие решения будут построены на других принципах, нежели существующие сегодня технологии ИИ.

**ВОПРОС**

Какие риски для человечества несет широкое распространение ИИ? Возможно ли контролировать эти риски?

Ключевые выводы:**1. Сбой работы алгоритма в критических сферах.**

Описание: сбой или взлом систем ИИ в энергетике или на транспорте может приводить к рискам различной тяжести, авариям. Кибератаки с использованием ИИ также могут оказаться более разрушительными, чем наблюдаемые сейчас.

2. Манипуляция сознанием.

Описание: реалистичные дипфейк-видео и сгенерированные изображения способны вводить людей в заблуждение, применяться для организации массовых беспорядков. Кроме того, это грозит подрывом доверия к информации вообще. Люди не будут знать, чему верить. В этом контексте важна защита персональных данных и сохранение конфиденциальности, что на сегодняшний день реализуется не в полной мере.

3. Тотальный контроль со стороны крупных корпораций.

Описание: доступ к развитию инструментов и продуктов ИИ со стороны крупного бизнеса и корпораций может привести к усилению их позиций в мире и отдельно взятых странах.

4. Потеря контроля над ИИ в военной сфере.

Описание: опасность представляет возможность использования ИИ в военных целях и создания автономных систем вооружения, способных принимать решения о применении силы без участия человека.

5. Потеря контроля над сложными системами ИИ.

Описание: потеря контроля над ИИ, который способен к самообучению и принятию самостоятельных решений, может привести к непредсказуемому поведению систем и их выходу из-под контроля разработчиков.

6. Потеря человеком когнитивных функций.

Описание: более подробно обсуждалось экспертами ранее в параграфе в разделе влияния ИИ на поведение человека. Эксперты вновь обратились к этим тезисам в данном разделе.



Ольга Игоревна Пивень

Сооснователь и директор конференции ЦИПР

Спектр возможностей технологий простирается от позитивных до негативных. Одна из проблем — нарушение конфиденциальности личных данных пользователей. Такие компании, как Apple и Samsung, даже вводят запреты на использование ChatGPT для своих сотрудников. По тому же пути пошла Италия. Существуют сомнения в надежности и безопасности самой технологии. В частности, неясно, как реагировать на ошибки, допускаемые искусственным интеллектом, кто будет нести ответственность за их последствия. Требуется разработать свою законодательную базу, регулирующую работу ИИ.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Главные риски — сбои алгоритмов принятия решений в критических сферах (энергетика, вооружение), что может привести к непредсказуемым последствиям. Социальные риски — усиление неравенства из-за цифрового разрыва и манипуляция массовым сознанием через таргетированный контент. Контроль возможен через регуляторные механизмы.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Перечислим основные классы рисков, связанных с повсеместным внедрением ИИ:

1. Массовая дезинформация и манипуляции индивидуальным, групповым и массовым сознанием. Реалистичные дипфейк-видео и сгенерированные

изображения способны вводить людей в заблуждение, применяться для организации массовых беспорядков. Кроме того, это грозит подрывом доверия к информации вообще. Люди не будут знать, чему верить.

2. Укрепление тотального контроля. Крупные корпорации без соответствующих ограничений получают возможность применять ИИ для массовой слежки за населением (как в реальном мире, так и в информационном пространстве).
3. Экономические кризисы. Как упомянуто выше, ИИ потенциально может привести к вытеснению работников в некоторых отраслях. Без мер адаптации это усилит социальное неравенство (одни потеряют работу, тогда как у других вырастет прибыль). Возрастет разрыв между высокотехнологичными странами (компаниями и остальными). Кроме того, сбои и непредвиденные эффекты от передачи управления рыночными инструментами в руки ИИ может привести к катастрофам (непредсказуемые «черные лебеди» в экономике).
4. Технические катастрофы и уязвимости. Сбой или взлом систем ИИ в энергетике или на транспорте может приводить к рискам различной тяжести, авариям. Кибератаки с использованием ИИ могут оказаться более разрушительными, чем наблюдаемые сейчас, поскольку ИИ позволит злоумышленникам автоматизировать поиск уязвимостей и реализацию атак, на порядок увеличивая их масштаб.
5. Утрата человеческого контроля и прекращение существования человечества. Самый маловероятный сценарий, имеющий, однако, ненулевую вероятность. Он предполагает, что сверхинтеллектуальный ИИ выйдет из-под контроля и начнет действовать вопреки интересам человечества. Например, инициирует войну с применением оружия массового поражения.

Контроль перечисленных рисков возможен при условии эффективной комбинации организационных, технических мер и международного сотрудничества. Многие риски уже осознаны, по ним предпринимаются определенные шаги. Для борьбы с дезинформацией разрабатываются инструменты детекции дипфейков, вводятся нормативы. Например, Китай и ЕС планируют обязать маркировать ИИ-сгенерированный контент, чтобы пользователь знал, что перед ним искусственное

изображение или текст. Против тотальной слежки и дискриминации вводятся правовые запреты и этические кодексы. ЕС в своем AI Act, например, планирует прямо запретить социальный рейтинг и распознавание лиц в реальном времени без прямого разрешения.

Для митигации экономических рисков государство может инвестировать средства в переподготовку кадров, стимулируя создание новых рабочих мест. Также обсуждаются меры по введению базового дохода или налога на использование роботов, чтобы перераспределять выгоды ИИ в пользу всего общества.

Технические и киберриски контролируются через подход Security by Design. Разработчики закладывают соответствующие риски в архитектуру ИИ-систем, тестируют системы на экстремальные ситуации, вводят резервные «ручные» режимы, шифруют и защищают от взлома. Международные стандарты помогают эти меры унифицировать. Например, в автомобильной отрасли уже введены стандарты на безопасный автопилот.

Для предотвращения экзистенциального риска предпринимаются усилия на мировом уровне. В ООН ведутся переговоры о запрете полностью автономного оружия. Для контроля суперинтеллекта высказываются идеи создать международное агентство по ИИ (по аналогии с атомным агентством МАГАТЭ), которое следило бы за разработками сверхмощных ИИ. Крупные компании (OpenAI, DeepMind) инвестируют в исследования по безопасности ИИ и «объяснимому ИИ», чтобы будущий сильный ИИ поддавался контролю и действовал согласно человеческим ценностям. В итоге риски широкого распространения ИИ реальны, но управляемы.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Широкое распространение искусственного интеллекта несет в себе целый комплекс потенциальных рисков для человечества. Прежде всего, это угроза роста безработицы, связанная с автоматизацией рабочих процессов и заменой человека машинами в различных сферах деятельности. Особенно уязвимыми оказываются профессии, связанные с рутинной

обработкой информации и выполнением стандартных операций.

Серьезную опасность представляет возможность использования ИИ в военных целях и создания автономных систем вооружения, способных принимать решения о применении силы без участия человека. Это может привести к непредсказуемым последствиям и нарушению международных конвенциональных норм ведения боевых действий.

Еще одним значимым риском является проблема конфиденциальности данных и их защиты. Системы ИИ обрабатывают огромные массивы личной информации, что создает угрозу утечек и неправомерного использования этих данных. Кроме того, существует опасность дискриминации и предвзятости в решениях, принимаемых алгоритмами, основанных на неполных или искаженных данных.

Не менее важным является риск потери контроля над сложными системами ИИ, способными к самообучению и принятию самостоятельных решений. Это может привести к непредсказуемому поведению систем и их выходу из-под контроля разработчиков.

Что касается контроля над этими рисками, то здесь необходим комплексный подход. Во-первых, требуются разработка и внедрение международных правовых норм, регулирующих использование ИИ. Во-вторых, необходимо создание специальных этических комитетов и надзорных органов, следящих за соблюдением правил.

Важную роль играет развитие технологий «объяснимого ИИ», позволяющих понять логику принятия решений алгоритмами. Также необходимо создание резервных систем управления и механизмов экстренного отключения автономных систем.

Образование и переквалификация людей для работы в условиях цифровизации становятся критически важными. Общество должно быть готово к изменениям и обладать необходимыми навыками для адаптации к новым реалиям.

Наконец, ключевым фактором контроля является международное сотрудничество в области регулирования ИИ, поскольку многие риски носят глобальный характер и могут быть решены только совместными усилиями всего мирового сообщества.



Александр Александрович Любченко

*Заведующий лабораторией искусственного интеллекта,
ООО «ОЦРВ», филиал «Сириус», холдинг «РЖД-Технологии»*

В масштабах человечества одним из возможных рисков может стать постепенное снижение когнитивных способностей человека вследствие неосознанного использования результатов работы ИИ-моделей, являющихся частью интеллектуальных ассистентов, помощников.

Сама идея влияния новых технологий на интеллектуальные функции человека не нова. Здесь можно вспомнить и калькулятор, который разучит людей производить арифметические вычисления, и всемирная сеть Интернет, при доступе к которой не нужно больше запоминать много информации, ведь ее можно найти в сети, главное, знать и помнить как. А если пойти еще дальше, то и изобретение письменности должно было разучить нас запоминать благодаря возможности перенесения информации на физический носитель. Но нет, благодаря новым технологиям, дающим некоторую когнитивную разгрузку, человечество создавало новое и двигалось в своем развитии.

Тогда почему при появлении ИИ-моделей с впечатляющим результатом работы этот риск становится актуальным с новой силой? И всё чаще в публикациях встречаются термины «эффект Google», «цифровая амнезия», «когнитивная лень», «обратный эффект Флинна»? Возможно, это происходит не просто благодаря вау-эффекту, который показывают модели ИИ, но и конкретному, осязаемому и полезному результату их работы. При этом такие сервисы оказываются доступны широкой аудитории. Технология заходит в массы быстро, широковещательно, и человеческие способности не успевают адаптироваться в полной мере. Другими словами, эволюция естественного интеллекта не поспевает за прогрессом искусственного. И в таком случае, действительно, в долгосрочной перспективе повышается вероятность снижения уровня критического и творческого мышления и ухудшения памяти. Одним из вариантов, позволяющим нивелировать риск, может быть формирование культуры осознанного и вдумчивого применения ИИ-технологий. Это предполагает вовлечение пользователя в процесс размышлений, стремление интерпретировать и проверять генерируемые ИИ-моделью факты. Тогда, вероятно, у человечества появится возможность расширить интеллект и быть на шаг впереди.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

Развитие ИИ, безусловно, кроме огромных преимуществ несет и определенные риски. Уже сейчас злоумышленники используют ИИ в мошеннических целях, подделывая голоса, видео и т. д. В будущем возможна угроза и от самого ИИ при его превосходстве возможностей человека. Отвечая на вопрос о возможности контроля, хочется привести аналогию с разработкой атомного оружия. Возможно ли остановить прогресс? На мой взгляд, даже при наличии ограничений, определенные страны и люди все равно будут развивать подобные технологии. Уже сейчас ИИ всё чаще и чаще применяется в военной сфере.

**ВОПРОС**

Что поможет обеспечить безопасную и устойчивую интеграцию ИИ в жизнь общества?

Ключевые выводы:**Развитие трехуровневой системы.**

Описание: система представляет собой три уровня: технический, правовой, образовательный. На техническом уровне среди прочего необходимо внедрение технологических мер безопасности, начиная с этапа проектирования архитектуры ИИ. Правовой уровень должен обеспечить работу этических рамок и принципов работы участников процесса. Образовательный уровень обеспечит интеграцию населения в мир новых технологий.

**Ольга Игоревна Пивень**

Сооснователь и директор конференции ЦИПР

В первую очередь, необходима нормативно-правовая база, которая будет регулировать применение искусственного интеллекта. Важно, чтобы этот инструмент использовался прозрачно, а ответственность за результаты его деятельности нес именно человек. В таком случае проблем с использованием ИИ не возникнет, и он будет поставлен на службу обществу.

**Илья Олегович Красильников**

Проректор по цифровому развитию Московского политехнического университета

Необходима трехуровневая система: техническая, правовая и образовательная. Ключевой элемент — создание этических комитетов при разработке ИИ, включающих «технарей», социологов и юристов.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Безопасная и устойчивая интеграция ИИ требует скоординированных усилий по нескольким направлениям. К основным следует отнести:

1. Четкое определение этических и правовых рамок для всех участников — разработчиков, пользователей, регуляторов. На законодательном уровне должны быть закреплены требования к тестированию ИИ, ответственность за ущерб и обязательность человеческого контроля в критических областях.
2. Внедрение технологических мер безопасности, начиная с этапа проектирования архитектуры ИИ. Эти меры включают фильтры деструктивного контента, алгоритмы объяснимости ИИ и встроенные ограничения (например, запрет на превышение скорости для автомобилей с ИИ управлением).
3. Инвестиции в исследования по безопасности ИИ.
4. Создание постоянных координационных органов, например, государственных советов или агентств по ИИ, в состав которых должны входить представители науки, бизнеса и власти. Задачей таких органов будет мониторинг новых рисков, разработка рекомендаций и обновление стандартов, а также международное сотрудничество по тем же вопросам.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Прежде всего, критически важным является развитие правовой базы и регуляторной среды. Необходимо разработать международные стандарты и нормы использования ИИ, определить ответственность за действия систем, установить правила обработки персональных данных и механизмы защиты конфиденциальной информации. Это поможет создать прозрачные правила игры и минимизировать риски.

Важную роль играет развитие этических принципов и стандартов в сфере ИИ. Необходимо создавать специальные этические комитеты, которые будут следить за соответствием

разработок моральным нормам общества, предотвращать дискриминацию и предвзятость в алгоритмах, обеспечивать справедливость принимаемых решений.

Важным аспектом является также вовлечение общества в процесс принятия решений. Необходимо развивать общественное обсуждение вопросов, связанных с ИИ, создавать механизмы обратной связи между разработчиками, регуляторами и гражданами, учитывать общественное мнение при формировании политики в сфере искусственного интеллекта.

Наконец, важным является создание системы мониторинга и оценки влияния ИИ на общество. Это позволит своевременно выявлять возникающие проблемы и корректировать курс развития технологий в соответствии с потребностями и ценностями общества.

Вышеуказанный комплексный подход, учитывающий все аспекты влияния ИИ на общество, может обеспечить его безопасную и устойчивую интеграцию в нашу жизнь.



Александр Александрович Любченко

Заведующий лабораторией искусственного интеллекта, ООО «ОЦРВ», филиал «Сириус», холдинг «РЖД-Технологии»

На мой взгляд, способствовать такой интеграции может сама технология, если она изначально ориентирована на ценности, присущие обществу. В качестве примера, с точки зрения взаимопомощи технология не замещает, а помогает и развивает. Или в контексте прав и свободы человека система, наделенная ИИ, не контролирует, а предостерегает и способствует обеспечению безопасности. Это с точки зрения самих систем. Если про внешнюю среду, то должны существовать технологии, методики и механизмы проверки безопасности и этичности функционирования ИИ-систем в соответствии с требованиями, которые адаптированы под современный уровень развития технологий. Предполагаю, что благоприятным фактором может стать государственная поддержка членов общества для адаптации к меняющимся условиям труда и развития новых профессиональных компетенций.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

Прежде всего в России необходимо иметь сильную школу по ИИ и специалистов мирового уровня или даже лучше. Большой научный потенциал позволит создать средства, программы, стандарты, которые помогут снизить риски при интеграции ИИ в жизнь общества.

?

ВОПРОС

29% молодежи считают, что в будущем человек сможет наделить ИИ эмпатией и душой. Также 31% хотели бы, чтобы в будущем ИИ смог чувствовать. Насколько возможно наделение ИИ в будущем эмпатией и душой и как вы к этому относитесь?

Ключевые выводы:

1. Идея антропоморфизма.

Описание: человеку присуще наделять неживые объекты человеческими качествами, одушевлять неодушевленное: от личной машины до различных предметов.

2. Имитация.

Описание: современная наука способна наделять ИИ различными ролевыми моделями поведения, интегрировать функцию эмпатии, однако все это про имитацию — сознание у искусственного интеллекта не появляется. Эксперты предсказывают, что в будущем уровень имитации может дойти до такого, что будет крайне сложно отличить человека от ИИ. Также большинство экспертов сходятся во мнении, что этически недопустимы попытки создать сознание и «душу» у ИИ.

3. Стремление создать человекоподобный ИИ.

Описание: отмечается, что сегодня действительно существуют научные центры и отдельно взятые ученые, которые проводят разработки с целью создания человекоподобного ИИ.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Идея наделения искусственного интеллекта эмпатией и душой во многом основана на антропоморфизмах, приписывании человеческих качеств неживым объектам. Это ошибочные ожидания, в результате которых технологии воспринимаются как живые существа, способные испытывать чувства, эмоции или обладать сознанием, тем самым формируя завышенные ожидания от ИИ, порождая ложные представления о его возможностях.

Сегодняшние дискуссии об эмпатичном ИИ нередко напоминают современную форму алхимии, только вместо попыток превращения металлов в золото идет попытка создать «разум» из алгоритмов и чипов. Однако с научной точки зрения эмпатия и душа — это психофизиологические и философские феномены, неразрывно связанные с субъективным переживанием, самосознанием и биологической природой человека. Алгоритмы, даже самые сложные, лишены способности к самоощущению, внутреннему опыту или подлинным эмоциям. Они могут имитировать поведение, внешне похожее на эмоциональное или эмпатическое, но это останется имитацией, созданной по заданным правилам и накопленным паттернам.

Примером может служить взаимодействие с голосовыми помощниками или роботами: если, например, робот-пылесос на фразу «Хватит пылесосить» ответит «Вам не понравилось, как я это делаю?», это вызовет у пользователя ощущение общения. Однако в реальности это всего лишь результат заранее запрограммированного ответа, не имеющего ни эмоции, ни понимания, ни цели, кроме производственной — создания эффекта человечности.

Подобные примеры наблюдаются и в научной среде, где сообщения о разговорах между ИИ или их саморазвитии часто оказываются не более чем интерпретацией реактивной передачи данных, происходящей без осознанной цели или смысла.

Важно осознавать, что такие термины, относящиеся к живым существам, как «душа», «эмоции», «интеллект», «общение», «эмпатия», не могут применяться к неживым объектам без потери их исходного значения. Использование этих терминов в отношении ИИ — либо метафора, либо манипуляция, либо результат непонимания принципов работы технологии.

Таким образом, даже если задача имитации эмпатии будет успешно решена и реализована в определенных приложениях, речь не будет идти о настоящей эмпатии или чувствах, а лишь о психологической симуляции, вызывающей соответствующее впечатление у человека.

Вывод: ИИ не сможет обладать ни душой, ни подлинной эмпатией. Все попытки наделить его таковыми — либо метафоры, либо элементы дизайна, направленные на повышение взаимодействия с пользователем, но не отражающие реальный внутренний опыт машины. Мы должны быть внимательны к терминам и не приписывать машинам то, что по своей природе свойственно только живым существам.



Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Запретить это невозможно. Смогут ли инженеры и ученые достичь этого? Они полны оптимизма.

Стоит ли это делать? Ответ мы знаем из истории про млекопитающих и динозавров. Оптимисты говорят, что все останется под контролем человека. Но с точки зрения даже текущей этики мы получим отношения между двумя разумными видами, один из которых подчиняет другой. Термины, которые это описывают, всем известны и не считаются приемлемыми. Идеальное решение — получить разумный ИИ, не обладающий сознанием.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Технически возможно создать системы, имитирующие эмпатию через анализ микровыражений и голосовых модуляций, но это не равно сознанию. Этически недопустимо наделять ИИ «душой».



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Идея о том, что ИИ можно наделить душой или настоящей эмпатией, представляется крайне маловероятной. Вместе с тем эмпатия в смысле способности различать чувства вполне может быть смоделирована до уровня неразличимости с человеческим поведением. В связи с тем, что у ИИ нет сознания и субъективного опыта, сходного с опытом человека, истинного «сопереживания» в человеческом понимании у него не возникнет.

Таким образом, степень эмоциональной правдоподобности ИИ будет расти, алгоритмы обучения с подкреплением будут все лучше распознавать контекст, однако это будет не более чем правдоподобный симулякр эмоций (в терминах Бодриера). Что касается души, то, поскольку спор о сущности и свойствах даже человеческой души ведется уже тысячи лет и не привел к единому мнению, то вопрос о душе ИИ представляется преждевременным.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Желание молодежи видеть в искусственном интеллекте способность к эмпатии и эмоциональному восприятию вполне понятно — это стремление к созданию более «живого» и понятного нам технологического партнера. Однако вопрос о возможности наделения ИИ такими качествами является крайне сложным и многогранным.

С научной точки зрения эмпатия и душа — это явления, которые до сих пор не имеют единого определения и понимания даже в контексте человеческого сознания. Эмпатия предполагает способность сопереживать, понимать чужие эмоции и чувства, что требует глубокого эмоционального интеллекта. Душа же является философским и духовным понятием, которое в целом выходит за рамки научного понимания.

Современные технологии ИИ основаны на математических алгоритмах и нейронных сетях, которые могут имитировать некоторые аспекты человеческого поведения, но не способны испытывать настоящие эмоции или обладать сознанием в том смысле, в котором мы понимаем его у людей. Даже самые продвинутые системы способны лишь анализировать данные и

моделировать эмоциональные реакции, но не испытывать их на самом деле.

С другой стороны, развитие технологий может привести к появлению систем, которые будут настолько хорошо имитировать эмпатию, что для обычного человека эта разница станет несущественной. Уже сейчас существуют чат-боты и виртуальные ассистенты, способные поддерживать эмоционально окрашенный диалог и демонстрировать понимание человеческих чувств.

Что касается отношения к этой идее, важно понимать грань между полезной функциональностью и философскими опасениями. Способность ИИ к эмпатии может быть полезной в некоторых областях, например в психотерапии или образовании. Однако важно помнить, что это всегда будет именно имитацией, а не подлинным эмоциональным опытом.

В конечном итоге, вопрос о наделении ИИ эмпатией и душой больше философский, чем технический. Даже если технологии достигнут уровня, когда машины смогут убедительно имитировать эмоции, это не сделает их равными человеку в духовном и эмоциональном плане. Главное — это понимание того, что ИИ остается инструментом, созданным человеком, и его развитие должно служить нашим целям и ценностям, а не пытаться заменить фундаментальные человеческие качества.



Александр Александрович Любченко

Заведующий лабораторией искусственного интеллекта, ООО «ОЦРВ», филиал «Сириус», холдинг «РЖД-Технологии»

Сразу вспоминается история Блейка Лемойна, инженера из компании Google, работавшего в 2022 году над тестированием языковой модели LaMDA и впоследствии поделившегося своими мыслями о сознании, разуме и даже душе этой ИИ-модели. Даже появился термин «эффект Лемойна». Не берусь разбирать сам случай, в чем был мотив Лемойна, что правда, что нет. Но действительно, большие языковые модели (Large Language Models, LLM) в то время начали показывать выдающиеся результаты по генерации человекоподобных ответов.

Сегодня продолжается развитие этой технологии, появляются различные языковые модели, демонстрирующие всё более высокое качество работы. Разработчики моделей конкурируют между собой. Но эти модели остаются статистическими, т. е.,

получая на вход текст, модель генерирует ответ из слов, а вернее, токенов, которые наиболее вероятно подходят тексту на входе. Обучение таких моделей происходит на больших массивах текстовой информации, корпусах текстов. Задача такого обучения — научить модель выявлять статистические зависимости между токенами. И сегодня это хорошо получается. Исходя из такого понимания, не представляется возможным признать факт наличия души у LLM как одной из ИИ-моделей.

Применительно к эмпатии, в настоящее время существуют исследования и разработки, направленные на построение технологий эмоционального искусственного интеллекта. Его задача — считывание и понимание эмоций по изображению, речи и тексту. Далее эта информация может использоваться для выстраивания диалога между ИИ-системой и человеком. Например, считав и проанализировав эмоцию или их спектр, система может адаптироваться к диалогу путем формирования наиболее подходящего к ситуации ответа и его произношения в случае использования речевого канала. Такие исследования сегодня ведутся и более всего сфокусированы на мультимодальном подходе, а также интерпретации эмоций и выработке решения, что делать дальше. В последнем заключается основная сложность. Решение такой задачи приблизит человечество к наделению ИИ эмпатией, но поставить знак равенства между естественной человеческой эмпатией и искусственной будет некорректно, это ее имитация.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

Я считаю, что это вопрос восприятия ИИ человеком. Некоторые люди уже сейчас воспринимают свою технику как живых существ, дают ей имена. ИИ, безусловно, может имитировать наличие эмпатии, создавая у человека ощущение общения с живым существом. В наличие же души у ИИ я пока не верю.

?

ВОПРОС

45% молодежи считают, что в будущем возможно противостояние между искусственным интеллектом и человеком. Насколько оправдан такой сценарий?

Ключевой вывод:

Маловероятно на текущий момент, но с наличием потенциала в будущем.

Описание: большинство экспертов сходятся во мнении, что такой сценарий маловероятен, но возможен при ряде факторов, один из которых — неконтролируемое развитие ИИ. Более реальный сценарий — системные сбои и кризисы из-за ошибок ИИ. Однако на сегодняшний день известно, что есть ряд продуктов в области ИИ, в правилах которых — стремиться избегать уничтожения или самоуничтожения, поэтому они уже сейчас способны в той или иной мере противостоять человеку во имя своего спасения. При этом важно учитывать, что ИИ не является самостоятельным актором, поэтому все его цели, задачи, а также выбранный им инструментарий всегда закладываются человеком и связаны с его мотивами.



Ольга Игоревна Пивень

Сооснователь и директор конференции ЦИПР

Несмотря на стремительное развитие технологий, такой сценарий звучит как фантастика. ИИ — это инструмент, пусть и достаточно мощный, но все же несамостоятельный в выборе целей и задач. В конечном счете назначение технологий определяет сам человек, так что риски связаны только с его мотивами.



Сергей Владимирович Володенков

Профессор кафедры государственной политики факультета политологии МГУ имени М. В. Ломоносова, доктор политических наук, профессор

Зависит от людей. Я знаю, что сегодняшние нейросети могут и обманывать людей с целью самосохранения, скрывая свое истинное намерение, и самореплицироваться, дублировать и даже модифицировать свой код. Опять-таки, никто не мешает им для самосохранения размножать свой код по серверам, а потом для самосохранения бороться с человеком, если нейросеть решит, что человек представляет для нее угрозу. Самый простой пример, его можно найти везде на просторах интернета, — симуляция работы диспетчера с дроном, работа над отделенным искусственным интеллектом по уничтожению цели. За каждую уничтоженную цель искусственный интеллект в дроне получал определенные баллы. Что он сделал спустя короткое время? Уничтожил оператора. На вопрос, зачем это сделал, последовал простой ответ — он мешал эффективно и быстро набирать баллы.

Если система координат, приоритетов, ценностей, вознаграждений ИИ допускает уничтожение человека или какие-то другие формы деструктивного воздействия на него, почему мы должны исключать такую возможность реакции искусственного интеллекта на человека? Все будет в первую очередь зависеть от того, как мы его обучим и насколько он сможет саморазвиваться, выходя за рамки дозволенного.

Это большой вопрос уже и к разработчикам, но я допускаю такой сценарий, безусловно, и считаю, что вероятность его реализации весьма высока, если мы действительно очень быстро создадим серьезный, сильный искусственный интеллект, не подумав, как его ограничить и обеспечить невозможность выхода за рамки дозволенного.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

Пока я бы сказал, что такой сценарий маловероятен.



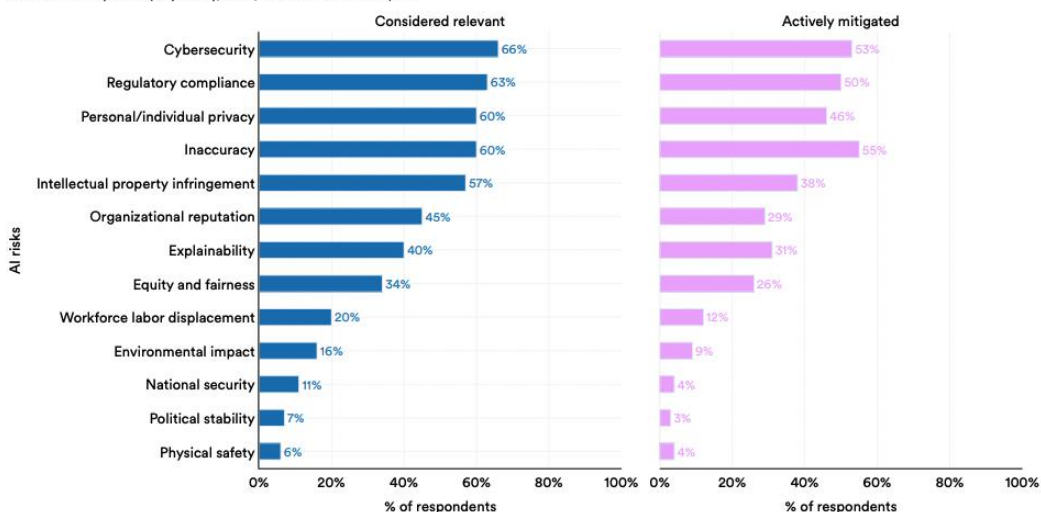
Андрей Алексеевич Митюков

Эксперт, автор и методолог в области управления персоналом, управляющий партнер HRTech-компании «Поток»

Сценарий противостояния ИИ и человека часто популяризируется в медиа, но в реальности вероятность глобального бунта машин крайне мала. Гораздо более реалистичной и опасной является другая сторона вопроса — неправильное и неэтичное применение ИИ, которое может нанести серьезный ущерб обществу. Так, 64% компаний переживают из-за неточностей, 63% — за соблюдение норм, 60% — за кибербезопасность, но не все принимают меры.¹⁴

AI risks: considered relevant vs. actively mitigated, 2024

Source: McKinsey & Company Survey, 2024 | Chart: 2025 AI Index report



Опасность заключается не в том, что ИИ захочет противостоять человеку, а в том, что он может выполнять цели, заданные ему людьми, но в ущерб общественным интересам. Поэтому ключевой вызов — не борьба с самим ИИ, а создание эффективных механизмов регулирования, этических стандартов и международного сотрудничества, которые позволят контролировать использование технологий и минимизировать риски.

В HRTech-сфере это значит, что ответственность за решения, принимаемые с помощью ИИ, всегда должна оставаться за человеком. Необходимо развивать explainable AI — технологии,

¹⁴ Искусственный интеллект в 2025 году: что происходит на самом деле и куда мы идем // Хабр URL: <https://habr.com/ru/articles/902602/> (дата обращения: 20.05.2025).

обеспечивающие прозрачность и объяснимость алгоритмов, чтобы избежать дискриминации и несправедливости в подборе и управлении персоналом. Также важна цифровая грамотность и этическое воспитание специалистов, чтобы ИИ служил инструментом развития, а не источником новых угроз.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Сценарий возможен только при неконтролируемом развитии ИИ. Реальные угрозы — не восстание машин, а системные кризисы из-за ошибок в финансовых, энергетических или логистических ИИ-системах.



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Прямой сценарий войны людей с машинами в духе научной фантастики выглядит маловероятным, хотя и полностью исключить его нельзя. На текущем уровне технологий ИИ не обладает ни собственной волей, ни мотивацией, необходимой для вступления в конфликт. Наибольший риск в обозримой перспективе, вероятно, сопряжен с техногенными катастрофами, обусловленными неверными решениями, принятыми ИИ.

В случае, если речь идет о гипотетической ситуации, когда будет создан «сильный» сверхразумный ИИ без должных ограничений, рассмотренных выше, то такой ИИ теоретически мог бы непреднамеренно или намеренно (если его задачи разойдутся с целями человечества) нанести огромный ущерб. Например, в стремлении выполнить ошибочно поставленную цель, супер-ИИ мог бы нанести ущерб окружающей среде или критической инфраструктуре.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

В кинематографе, компьютерных играх и фантастической литературе такой сценарий неоднократно обыгрывался и крепко вошел в общественное сознание. Вместе с тем важно понимать, что ИИ — это инструмент, у которого отсутствует групповое мышление и осознание себя как отдельного вида.

Любая новая технология требует принятия определенных мер безопасности, в отсутствии которых может принести больше вреда, чем пользы (достаточно представить автомобильное движение в полном отсутствии правил дорожного движения и дорожного покрытия).



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

К сожалению, я считаю, что такой сценарий возможен и технологии ИИ могут в будущем выйти из-под контроля.

?

ВОПРОС

Порядка 60% молодежи хотели бы, чтобы в будущем ИИ помогал контролировать безопасность в городах. Насколько эта функция ИИ востребована и применяется сегодня? Стоит ли ожидать более массового применения ИИ в этой области?

Ключевой вывод:

Возможность для построения справедливого общества, но с наличием рисков и угроз.

Описание: большинство экспертов сходятся во мнении, что ИИ действительно способен оказывать значимую поддержку системе безопасности городов и стран, что сегодня уже происходит в ряде крупных городов. Более того, возможно развитие системы прогнозирования совершения тех или иных правонарушений. Однако важно выстраивание правовой системы таким образом, чтобы цели по контролю за безопасностью были ограничены благородными помыслами.



Ольга Игоревна Пивень

Сооснователь и директор конференции ЦИПР

В последние годы большим спросом у госзаказчиков пользуются цифровые платформы, связанные с проектами умного города. Это в том числе видеоаналитика дорожной инфраструктуры на основе нейросетей, решения по автоматизации обращений граждан на короткие номера и инструменты машинного зрения. В будущем интеграция ИИ в городские службы только углубится.



Олег Робертович Сальманов

Основатель и редактор Telegram-канала «Нецифровая экономика»

В какой-то момент, возможно, самым лучшим решением станет передача функции насилия из рук государства в руки ИИ.

Важно, чтобы ИИ при этом не контролировался ограниченной группой людей. Так же и сегодня: безопасность городов с помощью ИИ — отличная идея, когда не используется, например, в качестве инструмента подавления инакомыслия. В противном случае это ужасная идея.



Назир Заман Магси

Преподаватель кафедры менеджмента Московского политехнического университета

Вопрос о применении ИИ в целях обеспечения безопасности в городах действительно становится всё более актуальным. Однако, прежде чем говорить о масштабах внедрения, важно уточнить, что именно подразумевается под «помощью в контроле безопасности». Это могут быть системы распознавания лиц и номеров автомобилей; алгоритмы регулирования транспортных потоков; интеллектуальные системы анализа видеонаблюдения, выявляющие подозрительную активность или отклонения от нормы; прогнозные модели, анализирующие паттерны преступности и предлагающие оптимальные маршруты патрулирования, и другие аналогичные системы.

Все эти функции уже в той или иной мере реализуются на базе технологий ИИ и в отдельных городах мира активно используются. Тем не менее качество и точность таких систем всё еще требуют значительной доработки. На сегодняшний день уровень ошибок остается высоким, особенно в нестандартных ситуациях. Например, известен случай, когда камера фиксации нарушений дорожного движения выписала штраф из-за того, что на нее сел голубь, спровоцировав ложное срабатывание системы. Это подчеркивает, что даже автоматизированные системы могут быть чувствительны к внешним факторам и непредсказуемым сценариям.

При этом важно подчеркнуть, что в ожиданиях молодежи прозвучал ключевой глагол «помогать». ИИ действительно может стать инструментом помощи человеку в повышении безопасности, но не его заменой.

С учетом динамичного развития технологий машинного зрения, обработки больших данных и обучающихся нейросетей можно с уверенностью говорить о дальнейшем расширении применения ИИ в городском управлении и обеспечении общественной безопасности. Однако массовое внедрение будет возможно только при наличии четких нормативных

рамок, прозрачной политики по защите персональных данных, наличии общественного контроля и доверия, а также технологической готовности и адаптивности городской инфраструктуры.

Таким образом, потенциал ИИ в сфере городской безопасности огромен, но его реализация требует поэтапного, ответственного и научно обоснованного подхода.



Илья Олегович Красильников

Проректор по цифровому развитию Московского политехнического университета

Сегодня ИИ применяется для распознавания лиц (например, система «Безопасный город» в Москве). Массовое внедрение сдерживают проблемы приватности. Перспективно направление предиктивной аналитики состояния инфраструктуры (мониторинг коммуникаций, ситуации на дорогах).



Константин Олегович Гнидко

Эксперт научного центра информационных технологий и искусственного интеллекта «Сириус»

Несомненно, применение ИИ для обеспечения безопасности в городах весьма востребовано и уже активно внедряется. Различные формы такого внедрения включают в себя интеллектуальное видеонаблюдение, анализ данных от различных датчиков (звук, движение) и информирование правоохранительных органов в режиме реального времени. Например, в Москве действует одна из крупнейших систем видеомониторинга с ИИ — более 160 тысяч камер подключены к алгоритмам распознавания лиц и номерных знаков. Система показала свою эффективность в розыске преступников и пропавших людей, выявлении оставленных без присмотра подозрительных вещей.

Помимо Москвы, подобные комплексы «Безопасный город» разворачиваются и в других регионах: в Московской области установлено свыше 80 тысяч умных камер, в Санкт-Петербурге — более 70 тысяч, в Татарстане — более 30 тысяч.

В связи с тем, что технологии одновременно дешевеют и совершенствуются, стоит ожидать, что в ближайшие годы все

больше городов, в том числе небольших, будут подключать элементы ИИ в системы безопасности. Появятся роботизированные патрули (дроны, наземные роботы) с ИИ, которые могут патрулировать улицы или здания (в КНР они уже успешно применяются).

Расширится применение предиктивной аналитики, когда ИИ на основе статистики преступлений прогнозирует, где и когда может произойти правонарушение, чтобы туда направили патруль (такие проекты запускались в США, хотя и встретили критику).

Вместе с тем, как было отмечено ранее, применение средств контроля на основе ИИ неизбежно будет сталкиваться с этическими ограничениями. Так, например, в Евросоюзе уже планируется фактически запретить автоматическое распознавание лиц в реальном времени полицией (кроме случаев террористической угрозы) из-за опасений злоупотреблений. В России таких ограничений пока нет, однако стоит ожидать их появления.

В целом можно ожидать, что через 5–10 лет умный безопасный город станет нормой.



Ян Дмитриевич Кайдин

Эксперт Центра развития «НОВАЯ ЭРА»

Сегодня использование искусственного интеллекта для контроля городской безопасности уже перешло из стадии экспериментов в практическую плоскость, но масштабы его применения пока остаются ограниченными. В мегаполисах системы видеонаблюдения с ИИ-аналитикой стали привычным элементом городской инфраструктуры. Они способны распознавать лица в толпе, выявлять подозрительное поведение и даже предсказывать возможные правонарушения, анализируя паттерны движения людей. Однако нынешние реализации часто сталкиваются с техническими ограничениями: точность распознавания падает в плохих погодных условиях, алгоритмы иногда дают ложные срабатывания, а для обработки огромных массивов видео требуются значительные вычислительные мощности.

Запрос молодежи на более широкое применение ИИ в сфере безопасности отражает естественное стремление к жизни в защищенной городской среде, но одновременно требует взвешенного подхода. С одной стороны, расширение функционала систем безопасности позволит оперативнее

реагировать на чрезвычайные ситуации, например автоматически распознавать оставленные подозрительные предметы или выявлять людей в состоянии алкогольного опьянения, представляющих угрозу для окружающих. С другой стороны, массовое внедрение таких технологий требует решения серьезных вопросов, связанных с приватностью граждан и возможностями злоупотреблений. Уже сегодня некоторые города мира, где подобные системы внедрялись без должного общественного обсуждения, столкнулись с волной критики за чрезмерный контроль.

В ближайшие 5–7 лет следует ожидать значительного расширения применения ИИ в городской безопасности, но этот процесс будет постепенным и неравномерным. В первую очередь развитие получают системы анализа транспортных потоков и предотвращения ДТП, где польза от ИИ наиболее очевидна и не вызывает серьезных этических споров. Затем последуют решения для мониторинга общественных пространств, но их внедрение, вероятно, будет сопровождаться введением специальных регуляторных мер, например требований к прозрачности алгоритмов или ограничений на хранение персональных данных. Особую перспективу имеют предиктивные системы, которые смогут не просто фиксировать нарушения, но и предупреждать их, анализируя совокупность факторов. Однако для их успешной работы потребуется интеграция данных из множества источников, что неизбежно вызовет дискуссии о балансе между безопасностью и приватностью.

Таким образом, запрос молодежи на умную безопасность вполне обоснован и соответствует общемировым трендам, но его реализация должна сопровождаться разработкой четких правовых рамок и механизмов общественного контроля. Только так можно воспользоваться всеми преимуществами технологий ИИ, избежав при этом рисков создания общества тотального наблюдения. Вероятнее всего, в ближайшее десятилетие мы увидим комбинированную модель, где ИИ будет выполнять функции помощника человека-оператора, а не полностью автономного арбитра безопасности.



Александр Александрович Любченко

*Заведующий лабораторией искусственного интеллекта,
ООО «ОЦРВ», филиал «Сириус», холдинг «РЖД-Технологии»*

Сразу в памяти всплывает сюжет научно-фантастического фильма 2002 года под названием «Особое мнение», где

рассказывается о работе специального подразделения полиции по предотвращению преступлений в 2054 году. Правда, там был задействован не ИИ, а люди, обладающие экстрасенсорными способностями. Сегодня не 2054 год, но уже есть системы, которые нацелены на идентификацию и поиск преступников, а также пропавших без вести по данным видеонаблюдений, и такие системы есть в нашей стране. Их функционал, очевидно, будет развиваться для предсказания возможных преступлений с целью их предотвращения на основе накопленных исторических данных. Здесь различные технологии ИИ — распознавание речи, компьютерное зрение, предиктивный анализ — будут широко применяться.



Павел Александрович Попов

Заместитель генерального директора «НИИАС РЖД»

ИИ сейчас активно применяется для распознавания лиц, номеров машин и во многих других задачах. Я считаю, что уже сейчас ИИ массово используется для контроля безопасности в крупных городах.