

ЭКСПЕРТНОЕ ЗАКЛЮЧЕНИЕ

на проект Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2030 года

В 2025 году истекло действие Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники на период до 2024 года, утвержденной распоряжением Правительства Российской Федерации от 19 августа 2020 г. № 2129-р. Ассоциацией «Альянс в сфере искусственного интеллекта» был подготовлен и направлен в Минцифры России проект Концепции развития регулирования отношений в сфере технологий искусственного интеллекта до 2030 года.

В связи с тем, что проект Концепции непосредственно затрагивает конституционные права и свободы человека и гражданина, Временной рабочей группой по защите прав и свобод человека и гражданина в цифровом пространстве Российской Федерации Совета при Президенте Российской Федерации по развитию гражданского общества и правам человека было подготовлено и утверждено настоящее экспертное заключение, которое будет направлено в Правительство Российской Федерации, а также опубликовано на сайте Совета.

I. Общая оценка проекта Концепции

- (1) Разработчикам проекта не удалось достичь сбалансированности при выборе подхода к правовому регулированию отношений в области искусственного интеллекта (далее – ИИ) и робототехники. В тексте содержится множество указаний на необходимость «мягкого регулирования», «саморегулирования», «повышения доверия населения к ИИ». Неоднократно повторяется тезис о развитии ИИ, которое не должно тормозиться регулированием, вводятся механизмы иммунитета для разработчиков ИИ, их защиты от юридической ответственности за работу систем ИИ. Обосновывается это в основном обещаниями будущих благ, которые должны появиться, когда такие системы проникнут во все сферы жизнедеятельности общества. Проект подготовлен бизнес-ассоциацией, участники которой стремятся к максимизации прибыли от внедрения технологий ИИ, однако государственная концепция не должна иметь настолько очевидный перекося в пользу коммерческих интересов частных компаний, пусть и с российской юрисдикцией.
- (2) Концепция предлагает один из возможных вариантов целеполагания в области государственного стратегического планирования. Следовательно, проект должен учитывать положения Федерального закона от 28 июня

2014 года № 172-ФЗ «О стратегическом планировании в Российской Федерации». В первую очередь, проект должен содержать обоснование предложенной цели развития ИИ в связи с основополагающими положениями стратегического планирования, принятыми Президентом Российской Федерации. Вместе с тем проект логически и содержательно не увязан с документами стратегического планирования. В частности, проект игнорирует ключевые положения Стратегии национальной безопасности Российской Федерации, утвержденной указом Президента Российской Федерации от 2 июля 2021 года № 400 «О Стратегии национальной безопасности Российской Федерации». Так, если в проекте заменить «Российская Федерация» на наименование любой другой страны вне зависимости от ее социально-экономического и геополитического положения, то принципиальных содержательных противоречий не возникнет. Последнее показывает, что предложенные в проекте идеи не основаны на учете культурных, социально-экономических и иных значимых национальных особенностей нашей страны. Далее, в тексте упоминаются «ценности», в т.ч. «охраняемые законом», но игнорируются традиционные российские духовно-нравственные ценности, приведенные в указе Президента Российской Федерации от 9 ноября 2022 года № 809 «Об утверждении Основ государственной политики по сохранению и укреплению традиционных российских духовно-нравственных ценностей». Наконец, проект должен отражать специальные, но не менее значимые отраслевые факторы, обозначенные в документах стратегического планирования Правительства Российской Федерации, например кадровый и научный потенциал, национальные потребности и ресурсы автоматизации и т.п.

- (3) Разработчики проекта некритически отнеслись к западным рейтингам и индексам, характеризующим степень внедрения ИИ в различные сферы экономики. В проекте не приведено разумного обоснования стремления к первым местам в таких рейтингах, вместо этого фактически содержатся призывы к участию в технологической гонке на чужих условиях. Также в условиях действия внешних неправомерных экономических ограничений весьма сомнительно выглядят тезисы о развитии экспорта технологий ИИ.
- (4) В проекте содержатся многочисленные и не всегда обоснованные ожидания повышения эффективности производства и качества продукции на основе только лишь внедрения технологий ИИ. В том числе, сообщается о предполагаемом повышении заработной платы во всех (!) отраслях в результате внедрения технологий ИИ. В качестве уже свершившегося достижения называется применение ИИ «в отраслях, традиционно

считающихся творческими и доступными только человеку», что является крайне спорным тезисом. Объективный подход требует отразить в тексте Концепции комплексную оценку перспектив внедрения ИИ, когда плюсы и минусы технологии могут компенсировать друг друга, а ее полезное использование потребует многолетних усилий.

- (5) В проекте крайне сжато и неполно описаны риски внедрения технологий ИИ, основной упор делается на риски безопасности и дискриминации. При этом не названы и не раскрыты уже реализовавшиеся риски массового мошенничества с использованием ИИ, применения ИИ для пропаганды и фальсификации, боевого применения ИИ, не обсуждаются проблемы ложных срабатываний технологий ИИ. Особенно важно подчеркнуть, что в проекте не упоминаются социальные риски внедрения технологий ИИ и последствиях излишней уверенности людей в качестве работы систем ИИ. Например, внедрение ИИ в некоторых отраслях национальной экономики, очевидно, обернется качественным изменением потребности в ручном низкоквалифицированном труде уязвимых групп населения, например инвалидов. Как будет распределяться бремя возникших социальных издержек в проекте не раскрывается и даже не ставится вопрос, о необходимости такого распределения. Очевидно, предполагается, что общество и государство примут все скрытые социальные издержки внедрения ИИ на себя.
- (6) В списке отраслевых направлений регулирования отсутствует военно-промышленный комплекс, хотя именно в этой сфере уже достигнуты значительные успехи в развитии технологий ИИ, а также, по оценкам специалистов, имеются все условия для создания прорывных технологий.

Таким образом, проект трудно назвать сбалансированным документом, в равной степени отражающим интересы личности, государства, бизнеса и общества, он не отражает реальной социально-экономической ситуации в нашей стране на текущем этапе, задачи сформулированы нечетко и недостаточно обоснованы, а риски представлены сжато и неполно.

II. Постатейные замечания к проекту Концепции

Глава I. «Общие положения», п. 1.

1. Требуется корректировка формулировки о том, что Россия «придерживается гибридного подхода» в регулировании ИИ. Такой тезис может быть использован в качестве направления дальнейшей деятельности, но не в качестве констатации.
2. Предлагается исключить упоминания о 4 месте России в рейтинге «комфортности регуляторной среды» и 7 месте в «Глобальном индексе присутствия», поскольку в проекте никак не обоснована необходимость борьбы за первые места в подобных западных рейтингах.
3. Приведённая позитивная оценка Федерального закона от 31 июля 2023 г. № 408-ФЗ (в части публикации наиболее важных параметров рекомендательных систем) не поддерживается специалистами на данном этапе его реализации. На практике имеет место крайне формальное исполнение цифровыми платформами данного законодательного требования.
4. Аналогичным образом обстоит ситуация с Кодексом этического регулирования в сфере ИИ и Декларацией об ответственной разработке и использовании сервисов в сфере генеративного ИИ. На практике многие положения Кодекса массово нарушаются в силу отсутствия каких-либо санкций. Это легко заметить по отсутствию предусмотренной Кодексом маркировки ИИ при взаимодействии с пользователями, которую не использует ни один крупный хозяйствующий субъект, заявивший о присоединении к Кодексу. Скорее, следует сделать акцент на необходимости соблюдения положений Кодекса его подписантами.
5. Вызывает сомнение формулировка о «ведущих экспортёрах российских технологий ИИ» в силу крайне скудного объёма экспорта в данной сфере.

Глава I. «Общие положения», п. 2.

6. Предлагается убрать термин «этического», т.к. совершенствование этического регулирования вряд ли может входить в компетенцию каких-либо органов публичной власти.

Глава I. «Общие положения», п. 3.

7. Задача «устранения излишних нормативных барьеров, ограничивающих развитие технологий ИИ» не подкреплена описанием или перечислением таких барьеров и характером их влияния на развитие технологий ИИ. Данная задача представляется надуманной.

8. Следует уточнить задачу о международном сотрудничестве «в целях продвижения национальных подходов в сфере ИИ, а также обмена лучшими практиками» указанием на то, что такое сотрудничество не может вестись с недружественными странами.

Глава I. «Общие положения», п. 4.

9. Из положения о недискриминации предлагается исключить слово «умышленной», т.к. неумышленная дискриминация обычно является следствием применения недоработанного (непроверенного) программного кода, т.е. аналогом выпуска некачественного товара на рынок.
10. Требуется уточнить положение об ответственности, т.к. из формулировки «ответственного нравственного выбора» не ясно, распространяется ли она, например, на случаи необоснованного отказа алгоритма ИИ в предоставлении банковского кредита или страхового продукта.
11. Требуется уточнить положение о квалифицированной экспертной оценке: во-первых, оно должно распространяться на регулирование всех случаев применения технологий ИИ (а не только их развития), во-вторых, в число специалистов следует также включить правоведов, социологов, психологов и представителей гражданского общества.
12. Из положения о запрете применения систем ИИ, «заведомо для разработчика приводящих к нарушению правовых норм», предлагается исключить слова «заведомо для разработчика».
13. Положение о «прозрачности и правдивости в информировании об уровне развития технологий ИИ, возможностях и рисках» необходимо уточнить, исключив, в частности, возможности ссылки разработчиков технологий ИИ на необходимость соблюдения режима коммерческой тайны и иные подобные лазейки.

Глава I. «Общие положения», п. 5.

14. Формулировка положения о необходимости «определения правового режима результатов интеллектуальной деятельности, созданных с помощью ИИ» предполагает необходимость закрепления гражданской правосубъектности ИИ, что не поддерживается большинством правоведов. Если охраноспособный результат интеллектуальной деятельности создан человеком при помощи технологии ИИ, его правовой режим определяется общими положениями об авторском праве. Сам ИИ не может заниматься творческой деятельностью в силу своей технической природы. Данную формулировку необходимо исключить. Также следует

исключить из п. 1 Главы II проекта указание на вопросы связанные с «защитой прав на контент, созданный с помощью технологий ИИ».

15. Из положения о маркировке и детекции контента, «сгенерированного технологией искусственным интеллекта» необходимо убрать опечатку.
16. Положение о том, что «факт использования систем ИИ не должен являться основанием для установления регуляторных ограничений» следует исключить, т.к. оно фактически исключает значительную часть регуляторного воздействия.

Глава II. «Общепромышленные задачи...», п. 1

17. Необходимо исключить манипулятивный и необоснованный прогноз резкого увеличения производительности труда и повышения оплаты труда специалистов во всех отраслях экономики. Напротив, следует говорить о рисках, связанных с массовым высвобождением работников вследствие автоматизации их рабочих мест, и подходах к их минимизации.

Глава II. «Общепромышленные задачи...», п. 2

18. Формулировки об антропоморфных роботах противоречат закреплённому выше в проекте Концепции запрету «антропоморфизации». Необходимо устранить противоречие.

Глава II. «Общепромышленные задачи...», п. 3

19. Задача по повышению доступности данных может быть достигнута на практике, прежде всего, за счет ослабления режима охраны различных их категорий, в т.ч., персональных данных. Остаётся неясным, кем поставлена такая задача и где она нормативно закреплена. Формулировку «адаптация законодательства» следует заменить на «соблюдение законодательства», имея в виду необходимость защиты прав граждан.
20. Формулировку о совершенствовании «контрольного просмотра» данных космической съемки необходимо доработать, исключив из под ее действия критически важные для безопасности и обороны объекты и территории.

Глава II. «Общепромышленные задачи...», п. 4

21. Положение о том, что выбранный законодателем подход к распределению ответственности не должен «способствовать созданию условий для незаконного привлечения к ответственности или для причинения вреда деловой репутации разработчиков и (или) пользователей и (или) владельцев систем ИИ, действия которых не привели к причинению вреда...», лишено правового смысла, т.к. причинение вреда является условием наступления ответственности (в том числе, таким условием

может быть причинение вреда деловой репутации). Данное положение предлагается исключить.

- 22.** В положении о необходимости «переноса бремени доказывания» не указано, в каких «отдельных случаях» это предлагается делать, а также на кого предлагается переносить это бремя. Не обосновано также, почему уход от ответственности должен способствовать «формированию доверия к ИИ со стороны общества». Данное положение предлагается исключить.

Глава II. «Общепромышленные задачи...», п. 5

- 23.** Утверждение о том, что во всем мире не выработан универсальный правовой подход к использованию результатов интеллектуальной деятельности для машинного обучения, весьма спорно, т.к. такое использование очевидно является случаем использования исключительных прав и должно осуществляться только с согласия правообладателя. Необходимость охраны исключительных прав вытекает из целого ряда нормативных актов, в т.ч., международных конвенций, ратифицированных нашей страной. Под «совершенствованием порядка» по-видимому, предполагается установление исключений для систем ИИ, что никак не обосновано и приведёт к появлению правовых коллизий. Данный пункт предлагается исключить целиком.

Глава II. «Общепромышленные задачи...», п. 6

- 24.** Среди задач в качестве самоочевидной заявлена задача повышения доверия граждан к технологиям ИИ. Эту цель необходимо исключить.
- 25.** Принцип доверенного ИИ необходимо распространить на все случаи его применения, а не только на те из них, где это применение может нанести ущерб безопасности нашей страны и правам граждан.
- 26.** Положение о необходимости «обеспечения доступа к качественным наборам данных» следует перенести в пункт 3, здесь же следует сохранить формулировку о необходимости развития критической инфраструктуры и защиты персональных граждан.

Глава II. «Общепромышленные задачи...», п. 7

- 27.** Утверждение о «высокой эффективности» механизма экспериментальных правовых режимов «в различных отраслях экономики» не подкреплено реальными доказательствами. Данное утверждение необходимо либо обосновать, либо исключить.
- 28.** Необходимо исправить опечатку в слове «совершенствования».

29. Также тезисы о страховании вреда, причинённого системами ИИ, следовало бы подтвердить данными об имеющейся статистике по количеству уже имеющихся случаях, величинах причинённого вреда и способах его возмещения. Статистика должна убедительно демонстрировать необходимость нормативных изменений, в противном случае они вряд ли нужны.

Глава II. «Общепромышленные задачи...», п. 8

30. Предлагается заменить систему добровольной оценки соответствия систем ИИ на обязательную в случаях, когда они применяются в критической информационной инфраструктуре, либо оказывают влияние на безопасность общества и государства, права и свободы личности.

31. Требуется дополнительное обоснование необходимости участия российских специалистов в разработке международных стандартов. Опыт такого участия в разработке стандартов по управлению Интернетом или в области информационной безопасности далеко не всегда конвертируется в реальное влияние на международные организации и проводимую ими политику. Фактическим результатом такого участия может стать внешнее управление российской индустрией ИИ.

Глава II. «Общепромышленные задачи...», п. 9

32. Задача по развитию механизмов саморегулирования в сфере ИИ уже ставилась ранее, поэтому на новом этапе должна быть предварена аудитом эффективности уже принятых решений. Как отмечалось выше, большинство подписантов документов этического саморегулирования используют их лишь в целях пиара и не во всем соблюдают на практике. Кроме того, конституционная задача по обеспечению безопасности личности, общества и государства при применении цифровых технологий (к которым относятся и технологии ИИ) должна реализовываться через нормативное регулирование.

Глава III. «Отраслевые направления...», п. 1

33. Вызывает обоснованные сомнения тезис о том, что «применение технологий ИИ в государственном (муниципальном) управлении прямо способствует улучшению доступа граждан к государственным услугам». На практике известны множественные примеры того, как цифровые технологии ухудшают доступ граждан к органам публичной власти, становясь дополнительным барьером при их взаимодействии. В этом пункте достаточно ограничиться перечислением имеющихся рисков.

Глава III. «Отраслевые направления...», п. 2

- 34.** Тезисы об обеспечении безопасности в школах и персонализации образования при помощи технологий ИИ не просто не обоснованы, но и опасны. Практика внедрения технологий ИИ в сферу школьной безопасности на практике часто выражается в нарушении личных прав обучающихся, их поголовном профилировании и цифровой слежке. Приоритетом развития системы образования должно стать формирование единого образовательного пространства на всей территории России, а не его фрагментация путём персонализации.
- 35.** Тезис о необходимости диагностики и анализа профилей учащегося «для формирования рекомендаций по возможным профессиональным траекториям» также не выдерживает никакой критики со стороны профессиональных педагогов и родительского сообщества. Этот тезис базируется на устаревших подходах педологии и утверждении о линейном и управляемом характере развития человеческой личности. Данный тезис не подлежит исправлению, его необходимо исключить.
- 36.** Аналогичным образом за тезисом о необходимости использования «систем прокторинга с ИИ» скрывается еще один инструмент цифровой слежки за обучающимися, что прямо запрещено Конституцией РФ и в других местах проекта Концепции. Этот тезис также необходимо исключить.

Глава IV. «Механизмы реализации Концепции»

- 37.** Вызывает обоснованные сомнения необходимость «совершенствования механизмов упрощённого внедрения ИИ», подразумевающая дальнейшее ослабление законодательных ограничений. С учётом описанных в проекте Концепции рисков следует напротив, рассмотреть вопросы совершенствования процедуры допуска технологий ИИ на рынок, включающие в необходимых случаях оценку соответствия, страхование вреда, тестирование на доверенность и аудит их влияния на права и свободы личности.