

Научная статья
УДК 34.09:004.8:351/354

ПЕРСПЕКТИВЫ ПРАВОВОГО РЕГУЛИРОВАНИЯ В СФЕРЕ СИЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ГОСУДАРСТВЕННОЕ УПРАВЛЕНИЕ: ПОПЫТКИ ОГРАНИЧЕНИЯ БЕЗГРАНИЧНОГО

Полстовалов О. В.

АНО «Содействие и развитие инноваций в научно-производственной сфере»,
г. Москва, Россия

Аннотация. В статье раскрываются основные подходы к этизации применения сильного искусственного интеллекта, по сути, как на заре становления, так и в современных условиях внедрения этой сквозной технологии. Представлен анализ стратегического целеполагания развития искусственного интеллекта, предлагается обобщение зарубежного и международного опыта по вопросу инициатив в сфере этических и правовых ограничений. Высказываются рекомендации по совершенствованию российской практики стратегирования, нормотворчества и этизации в области применения искусственного интеллекта. Обобщены тематические поручения Президента России, инициативы Консультативного органа высокого уровня ООН по искусственному интеллекту «Управление искусственным интеллектом в интересах человечества», принятые документы на уровне ЕС, подходы к этизации закрытого клуба «Альянс в сфере ИИ».

Ключевые слова: искусственный интеллект, стратегические, законодательные и этические подходы и ограничения, перспективы развития.

PROSPECTS FOR LEGAL REGULATION IN THE FIELD OF STRONG ARTIFICIAL INTELLIGENCE AND PUBLIC ADMINISTRATION: ATTEMPTS TO LIMIT THE LIMITLESS

O. V. Polstovalov,

ANO «Assistance and Development of Innovations in the Scientific and Industrial Sphere», Moscow, Russia

© Полстовалов О.В., 2026

Abstract. *This article explores the main approaches to the ethicalization of strong artificial intelligence, both in its infancy and in the current context of this end-to-end technology's implementation. It presents an analysis of the strategic goal-setting for the development of artificial intelligence and offers a summary of international and international experience on initiatives in the area of ethical and legal restrictions. Recommendations are made for improving Russian practices in strategizing, rule-making, and ethicalization in the field of artificial intelligence. The article summarizes the thematic instructions of the President of Russia, the initiatives of the UN High-Level Advisory Body on Artificial Intelligence "Governance of Artificial Intelligence for Humanity," documents adopted at the EU level, and approaches to the ethicalization of the closed club "AI Alliance."*

Keywords: *artificial intelligence, strategic, legislative, and ethical approaches and restrictions, development prospects.*

Правовое регулирование в сфере искусственного интеллекта во всем мире зарождается в условиях высокой неопределенности перспектив и пределов развития этой сквозной технологии. Даже отграничение искусственного интеллекта от сильной его версии носит весьма субъективный характер в опоре на возможности человека, которые, как известно, у разных людей весьма различные.

Целеполагание в Национальной стратегии развития искусственного интеллекта (далее иногда – ИИ, сильный искусственный интеллект – СИИ) на период до 2030 года разбивается на несколько составляющих, что свидетельствует о том, что главный ориентир здесь выделить либо невозможно, либо его многосоставность находится в органическом единстве всех элементов, которые корреляционно со всей мерой необходимости связаны между собой. Однако это не так. Более того, целевые ориентиры развития ИИ при неопределенности прогноза «к чему все это приведет вообще» и без разбивки по уровням на основе качественных характеристик при отсутствии общепризнанного понимания самой ключевой категории «искусственный интеллект» вообще ставит под сомнение возможность

конструирования общественно полезного результата, во-первых, непосредственно являющегося следствием процесса принимаемых при реализации стратегии усилий, во-вторых, опирающегося на прогноз по его достижимости, в-третьих, основанного на общепризнанной или, по меньшей мере, максимально адекватной действительности терминологии.

Логика «цель-задачи» всегда присутствует в любой рациональной деятельности. К достижению цели ведет решение частных задач, но проблема современного стратегирования заключается в том, что операционно не напрямую связанные результаты ставятся во главу угла. Порой в целеполагание попадает выборка отдельных прогнозно-неопределенных ориентиров. «Целями развития искусственного интеллекта в Российской Федерации, – указано в п. 23 Национальной стратегии развития искусственного интеллекта, – являются обеспечение роста благосостояния и качества жизни ее населения, обеспечение национальной безопасности и правопорядка, достижение устойчивой конкурентоспособности российской экономики, в том числе **лидирующих позиций в мире** (выделено – О. П.) в области искусственного интеллекта» [1]. В этом конгломерате из элементов «3+» сложно найти единый вектор общей логики и корреляции. При этом высокая степень неопределенности перспектив развития СИИ по качеству, времени и пространству не позволяет нам упрекнуть в нечеткости и разрозненности целей. Отметим только социальную ориентированность по приоритетам, поскольку благосостояние граждан находится на первом месте.

Любая цель развития как некий конечный пункт должна быть достижима. При этом состояние лидерства само по себе дает преимущества на определенную перспективу, а позиционирование такового «среди прочих других», в том числе геостратегических противников, не дает нам уверенности в обеспечении национальной безопасности. В любом случае

стратегическая цель как конечный пункт во временной перспективе не может быть навсегда и окончательно данным состоянием. Это динамическая система. Иными словами, нас могут догнать, если это лидерство не обеспечивать в постоянном режиме, или нас могут «вытолкнуть», если такое лидерство носит коллективный характер. Однако применительно к единоличному лидерству в сфере ИИ, если принимать за основу безграничность прикладного использования СИИ, все может быть совершенно не так. Главный вопрос, на который не решаются вслух ответить однозначно: достижение надежного сильного искусственного интеллекта будет означать не только мировое лидерство, но и научно-технологическую глобальную гегемонию, или все-таки возможности достижения «новой Атлантиды» или «новой Шамбалы» сильно преувеличены? ИИ дает результаты в строго формализованных системах, где есть в причинном комплексе корреляционно-определяющие составляющие, например, в системе профилактики смертности в ДТП. Способен ли даже СИИ справиться с задачей генерации нового литературного творчества – большой вопрос, поскольку синтез нескольких стилей, персонифицированных семантических графов (Е. В. Гарин) гениальных писателей и поэтов не порождает нового творца. Будущая литература от СИИ – век вечных пародий? СИИ установит непреодолимые пределы человеческого творчества?

В целеполагании о лидерстве должны быть более точные содержательные (что под этим конкретно понимается) и временные (когда это случится) ориентиры, которые невозможно определить без научно обоснованного достоверного прогноза, опирающегося на анализ конкурентных возможностей соперников. Президент России Владимир Путин еще в 2017 году отметил, что за развитием искусственного интеллекта стоит не только будущее России, но и всего мира, а тот, кто займет лидерские

позиции в этой сфере, получит и глобальное лидерство [2]. Множественное число «лидерские позиции» не исключает концепции «коллективного доминирования», где возможна, как и в любой отрасли, сегментация по отдельным направлениям. Думается, что это будет «доминирующая внутривидовая конкуренция», поскольку гегемоны вряд ли договорятся, и дело здесь не в геополитических симпатиях и антипатиях.

В глоссарии Национальной стратегии развития искусственного интеллекта на период до 2030 года искусственный интеллект понимается как комплекс «технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их». Иными словами, имитация, а не замещение, дает некую надежду на то, что контроль остается за человеком и, к примеру, неантропоцентричная наука не выглядит однозначной и непреодолимой перспективой будущего, но возможность СИИ превосходить по результатам интеллектуальную деятельность человека не оставляет людям конкурентных преимуществ за исключением, пожалуй, права на первоисточник. Более того, если признавать силу естественного отбора, то именно он направлен «на формирование и эволюционное развитие у организмов все более высокоорганизованных когнитивных систем, **способных информационно контролировать окружающую среду** (выделено – О. П.) и их собственные когнитивные состояния (самовосприятие) с помощью создаваемой этими системами когнитивной информации [3, с. 22]. Иными словами, возникает риторический вопрос: каковы параметры и масштабы эволюции искусственного интеллекта, где человеческое видится лишь на старте, а все остальное уходит в неконтролируемую людьми плоскость?

Все это задает требования, как минимум, к этизации процесса на основе компромиссов в начале исследований, но, увы, перспективы ИИ слишком заманчивы, чтобы остановить надвигающуюся турбулентность. Логика российской науки проста, как и рациональность политики: необходимо максимально ускориться в процессе разработки сильного искусственного интеллекта, дабы не отстать от мировых лидеров, не упустить инициативу и в лучшем случае не остаться на задворках мировой гегемонии. И эта всеобщая гордыня приведет к мировому разрушению антропоцентрии? Так пусть это случится при лидирующих позициях России? А зачем тогда вообще лидирующие позиции? Для мировой гегемонии в мире, где уже нет человека в его сегодняшнем понимании? Сплошные гипотезы. При таком уровне гипотетического формирования стратегии развития ИИ может превратиться в известный сюжет русской народной сказки «Пойди туда – не знаю куда, принеси то – не знаю что». Ограничительные меры, принимаемые в мире и его регионах даже на межгосударственном уровне, пока продиктованы предупредительной риск-ориентированной мотивацией. Генеральная ассамблея ООН в марте 2024 г. приняла первую резолюцию об искусственном интеллекте, представленную по инициативе США и состоящую из тринадцати пунктов, среди которых особенно «успокаивают» положения о необходимости преодоления разрывов в сфере ИИ и других цифровых технологий между странами и в них, о помощи развивающимся государствам для обеспечения равного и справедливого доступа к цифровой трансформации и безопасным системам ИИ и пр. [4].

Законодатели Европейского Союза достигли политического соглашения по проекту закона об искусственном интеллекте (ИИ) в декабре 2023 г., а утвердили его в марте 2024 г. Данный закон основан на риск-ориентированном подходе до уровня неприемлемости отдельных новаций и

направлений. Некоторые системы ИИ, представляющие «неприемлемые» риски, запрещены. К широкому спектру систем искусственного интеллекта «высокого риска», которые могут оказать пагубное влияние на жизнь людей, их здоровье, безопасность, основные права, применен условно-разрешенный режим: при соблюдении ряда требований и обязательств по получению доступа на рынок ЕС. Системы искусственного интеллекта создают риски из-за отсутствия у них прозрачности. Те из них, которые представляют лишь минимальный риск для людей, не будут подлежать ограничениям. Интересен был сам процесс обсуждения в декабре 2023 г. Исследование, проведенное по заказу Европейского парламента, рекомендовало, в частности, уполномочить профильную Комиссию периодически адаптировать список запрещенных практик ИИ под наблюдением Европарламента, а также принятие более полного перечня «ограниченных приложений ИИ», включающего удаленную биометрическую идентификацию в режиме реального времени без ограничений для правоохранительных целей. Регулирование технологии распознавания лиц (FRT) было одним из наиболее спорных вопросов. В частности, Европейский совет по защите данных (EDPB) призвал к общему запрету на любое использование ИИ для автоматического распознавания черт лица человека в общедоступных пространствах [5, р. 1, 5-6]. Дискуссия велась очень жесткая, однако в соответствии с согласованным документом было запрещено использование в режиме реального времени в общественных местах систем распознавания лиц и других систем «удаленной биометрической идентификации», как и распознавание эмоций и использование полицией предиктивной аналитики для профилактики правонарушений. Исключение сделано для некоторых конкретных случаев. В частности, полиция сможет использовать такие технологии при угрозе теракта, для поиска жертв или в рамках расследования тяжкого преступления

[6]. Убежден, что все эти нормы-ограничения – чистой воды «белый шум» и некие реверансы к давно ушедшим на периферию реальной политики западным ценностям прав человека, о которых нельзя не сказать и даже необходимо создать видимость важности их соблюдения. Искренне верующие в непоколебимость значимости прав человека актеры на этой сцене становятся все большей редкостью, но они необходимы для придания правдоподобности самому спектаклю с бездарной труппой функционеров ЕС. Трехдневный спектакль окончился устраивающими всех формулировками, где риски причинения вреда все-таки остались в сфере волеизъявления поставщиков услуг, читай – квалифицированных заказчиков.

Как известно, в марте 2024 г. закрытый клуб «Альянс в сфере ИИ» подписал Декларацию об ответственной разработке и использовании сервисов на основе генеративного искусственного интеллекта (далее – Декларация). Наконец-то мы видим, что негенеративный ИИ вряд ли заслуживает стратегирования ввиду отсутствия высоких рисков, связанных с его использованием. Действительно, кому может повредить уровень ИИ по типу «Т9»? «Потенциал внедрения генеративного искусственного интеллекта – отмечено в преамбульной части Декларации, – для решения конкретных задач безграничен» [7]. Интересно: как можно определить целеполагание безграничного или, уж тем более, его ограничить посредством такой вот «умеренной этизации»? Оставим за скобками не совсем удачные обороты «рекомендации ..., которыми рекомендуется руководствоваться» (п. 1 Декларации) с тем, чтобы акцентировать внимание на главном, на риск-ориентированной работе над документом. Поражает сразу первая фобия: «Обеспечивайте безопасность и надежность ваших сервисов на основе генеративного искусственного интеллекта. Принимайте меры для предотвращения злоупотреблений и создания некорректной информации».

Несложно себе представить красные стрелочки с указанием «alarm». Еще один риторический вопрос: как обеспечить безопасность от того, чей потенциал внедрения безграничен? К некорректной информации относятся невежливые, неучтивые, попросту грубые сообщения. Или речь идет только о достоверной информации, и о какой степени соответствия действительности может здесь идти речь? Есть ли здесь место научным гипотезам вообще, об истинности или ложности которых рано судить до завершения исследования? Судя по п. 4 Декларации, некорректная информация – это есть нечто отдельное от информации незаконной.

Пятый пункт в рекомендациях для разработчиков Декларации гласит: «Информируйте пользователя о факте использования генеративного искусственного интеллекта в ваших сервисах там, где это возможно и неочевидно». Успокаивает диспозитивное «там, где это возможно». В ЕС дипфейки, согласно приведенному выше Закону-регламенту, подлежат маркировке. Иными словами, генеративный ИИ, искажающий информацию до степени максимальной достоверности, подлежит, как минимум, маркировке. Что делать с розыгрышами посредством дипфейков в России?

В другом этизирующем разработку и применение ИИ документе – Кодексе этики в сфере искусственного интеллекта [8] (далее – Кодекс), мы наблюдаем абсолютно нестандартный подход как по форме, так и по содержанию с точки зрения основ юридической техники. Разумеется, это не кодекс в полном смысле этого слова, поскольку он ничего не кодифицирует, да и его объем говорит больше о формате декларации о намерениях, о том, насколько предстоит вообще разобраться с перспективами ИИ: куда эта тропинка заведет человечество в целом и российское общество в частности. Свои опасения авторы Кодекса раскрывают сразу, не откладывая дело в долгий ящик: «Актеры ИИ должны принимать необходимые меры,

направленные на сохранение автономии и свободы воли человека в принятии им решений, права выбора и в целом сохранения интеллектуальных способностей человека как самостоятельной ценности и системообразующего фактора современной цивилизации» (п. 1.2). Самое нелепое, но, впрочем, подкупающее своей необязательностью правило заключается в том, что на этапе создания СИИ возникает обязанность акторов ИИ по прогнозированию возможных негативных последствий для развития когнитивных способностей человека, помимо того они не должны допускать разработку СИИ, которая целенаправленно вызовет такие последствия. Проблема в том, что любой условный гаджет не делает нас умнее, о чем свидетельствуют наблюдения за деградацией интеллектуальных способностей населения в их прямой пропорциональной зависимости от роста средств и возможностей по передаче на аутсорсинг цифровым сервисам мыслительной деятельности человека. Любая оптимизация поиска интересующей информации, безусловно, упрощает форму умственного труда, но значительно опаснее, когда мы получим превосходящие познавательные человеческие возможности ресурсы СИИ, которые выступят вполне себе достойной альтернативой. Перспектива ремонтно-восстановительного стада для человечества в этой связи уже не только не выглядит утопией, но, скорее всего, не является худшим из возможных сценариев развития событий. И здесь высоки риски превращения человека из субъекта-творца в обленившегося и вымирающего в интеллектуальном поле пользователя-объекта, который даже не замечает метаморфозы его ключевой роли. Интеллектуальная и творческая депопуляция элиты общества в этой связи представляется высоко вероятной. Элитаризация и популяризация людей недалекого ума уже никого не удивляет, да и условия к тому сложились как никогда кстати. Впору рассуждать об «искусственном антиинтеллекте».

В любом случае стратегия должна, на мой взгляд, отвечать следующим основным требованиям:

1. Целеполагание может быть определено только через мировое единоличное лидерство (пусть даже в отдельных направлениях), но научно спрогнозированное и взвешенное в международном масштабе. Иное целеполагание в рамках принятого документа будет выглядеть не более чем «белым шумом». Представляется, что такое целеполагание и часть стратегии должны уйти в формат «закрытых» нормативно-правовых актов. Технологическое лидерство в сфере ИИ определено пп. «а» п. 7 Указа Президента Российской Федерации от 7 мая 2024 г. № 309 «О национальных целях развития Российской Федерации на период до 2030 года и на перспективу до 2036 года» (далее – Указ-309). Само по себе технологическое лидерство представляет собой «превосходство технологий и (или) продукции по основным параметрам (функциональным, техническим, стоимостным) над зарубежными аналогами» (см.: раздел II. «Основные понятия» Концепции технологического развития на период до 2030 года, утв. Распоряжением Правительства РФ от 20 мая 2023 г. № 1315-р).

2. Этизация разработки и пользования продуктами ИИ в действующих вариантах никого и ни к чему не обязывает, поэтому заседания различных высокотехнологичных клубов по интересам лишь показывают невозможность ограничения того, что, действительно, имеет безграничный ресурс по внедрению. Обязывать и ограничивать в общественных интересах может лишь локализация самого процесса в специальной госкорпорации с абсолютно публичной природой и без права на приватизацию.

3. Необходимо обеспечение чистоты и достоверности исходных данных для применения ИИ, которые не достигнуты, к примеру, для решения задачи формулирования и внедрения на основе искусственного интеллекта

инициатив по разработке стратегии пространственного развития России ввиду отсутствия адекватных действительности сведений даже в формате экономической географии нашей страны. Более того, все происходящие изменения с управляемой системой (базовые значения тоже меняются) подлежат строгому учету для корректировки выбранного курса. Поэтому детальная и максимально объективная инвентаризация кадровых, природных, инфраструктурных ресурсов, имеющихся научных школ и достижений в формате, к примеру, закрытого реестра открытий (предложение Е. В. Гарина) просто жизненно необходима. Меняются и внешние условия, поэтому исходные базовые значения жизненно необходимы на старте, но события, явления и факты объективной действительности развиваются в пространстве и времени, в условиях реальных исторических событий, общественно-политических метаморфоз, глобальных и региональных конфликтов. Основоположник отечественной экономической кибернетики Н. И. Ведута справедливо отмечал, что сведения о событиях нас интересуют не из любопытства, а ввиду необходимости управлять их ходом: «В ряде случаев можно располагать всей информацией, необходимой для принятия точного решения об управляющем воздействии, обеспечивающем перевод системы в требуемое состояние. Однако это достигается в пределах вполне определенной точности замеров. Такая точность не беспредельна, и в связи с неизбежным накоплением ошибок для принятия последующих решений необходимо получение новой информации о состоянии системы» [9, с. 59]. Добавим, что в современной системе международных отношений и «обмена научным опытом» с иностранными коллегами, помимо ошибок, необходимо считаться и со ставшей практически нормой если не прямой, то скрытой дезинформацией, с откровенной и завуалированной ложью, неточностью

интерпретаций и политическим новоязом, в основе которого органично помещается сознательное искажение действительности в эгоистических целях.

Следует остановиться на поручениях Президента Российской Федерации и действующем отечественном законодательстве, в рамках которых определились контуры работы с тем, что мы понимаем под СИИ. С 2024 года тема искусственного интеллекта становится мейнстримом в повестке обсуждения Президента и Правительства, заседаний правоохранительных ведомств.

Президент России поручил Правительству «при формировании федерального проекта «Искусственный интеллект» в рамках разработки национального проекта «Экономика данных» предусмотреть меры поддержки развития и внедрения больших генеративных моделей и технологических решений в области искусственного интеллекта, а также создание инфраструктуры для их широкого использования» (Пр-83, п.1-в от 17.01.2024). Технологии искусственного интеллекта в отраслях экономики, социальной сферы (включая сферу общественной безопасности) и в органах публичной власти вошли в перечень сквозных важнейших технологий согласно п. 25 Указа Президента Российской Федерации от 18 июня 2024 г. № 529 «Об утверждении приоритетных направлений научно-технологического развития и перечня важнейших наукоемких технологий» (далее Указ-529). И в этом же Указе Главы государства Правительству Российской Федерации было поручено в трехмесячный срок «определить объем финансирования и порядок реализации мероприятий, связанных с разработкой и ускоренным внедрением в экономику важнейших наукоемких технологий, в том числе в рамках национальных проектов по обеспечению технологического лидерства». Не случайно в число целевых показателей и задач, выполнение которых характеризует достижение национальной цели «Технологическое

лидерство», вошло обеспечение технологической независимости и формирование новых рынков по таким направлениям, как экономика данных и цифровая трансформация, искусственный интеллект (пп. «а» п. 7 Указа-309). Национальные проекты технологического лидерства должны быть ориентированы на достижение превосходства технологий и (или) продукции по основным параметрам (функциональным, техническим, стоимостным) над зарубежными аналогами. Именно такое понимание технологического лидерства заложено в Концепции технологического развития на период до 2030 г., утв. Распоряжением Правительства РФ от 20 мая 2023 г. № 1315-р.

Первое поручение Президента, которое было исполнено в вышеназванном формате, звучало так: «Минэкономразвития России совместно с ассоциацией «Альянс в сфере искусственного интеллекта» и другими заинтересованными организациями обеспечить координацию работы по подготовке в рамках Кодекса этики в сфере искусственного интеллекта декларации об ответственной разработке и использовании больших генеративных моделей» (Пр-83, п. 8 от 17.01.2024). Несмотря на качество документа, это был все-таки определенный шаг к этизации в сфере искусственного интеллекта.

Куда как более жесткой была формулировка другого поручения Главы государства, где Правительству Российской Федерации совместно с Государственной Думой Федерального Собрания Российской Федерации было поручено «обеспечить внесение в законодательство Российской Федерации изменений, направленных на упрощение доступа хозяйствующих субъектов к участию в экспериментальных правовых режимах в сфере цифровых инноваций и установление ответственности за причинение вреда жизни, здоровью и имуществу граждан в результате тестирования и использования технологий искусственного интеллекта» (Пр-83, п. 2 от

17.01.2024). Срок исполнения был определен 1 февраля 2024 г. Однако поручение было исполнено в части установления ответственности, да и то в рамках экспериментального правового режима, да еще и с отлагательным сроком вступления данных изменений в действующее законодательство спустя 180 дней с момента опубликования закона. Иными словами, через сто восемьдесят дней с 08.07.2024. Не вступившим еще в силу законом было оговорено, что «в случае, если при реализации экспериментального правового режима в результате использования решений, разработанных с применением технологий искусственного интеллекта, причинен вред жизни, здоровью или имуществу человека либо имуществу юридического лица, в течение тридцати дней со дня выявления факта причинения такого вреда регулирующий орган создает комиссию для установления обстоятельств, при которых был причинен такой вред» (ч. 1 ст. 18.1 Федерального закона от 31.07.2020 № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации»). Таким образом, отчасти особенности юридической ответственности за решения на основе ИИ, которыми был причинен вред в рамках экспериментального правового режима, находят свое, пока еще робкое, отражение в законодательстве. Остается открытым вопрос о том, насколько данное положение будет гармонизировано с нормами гражданского законодательства в части возмещения причиненного вреда. В рамках разработки национального проекта «Экономика данных» Правительству РФ было поручено Президентом России предусмотреть «повышение доли предоставляемых массовых социально значимых государственных и муниципальных услуг в электронном виде в проактивном режиме с использованием, в том числе, технологий искусственного интеллекта в общем объеме таких услуг». По большому счету, увеличение доли оказания государственных услуг цифрового формата является

нормальной практикой, и поэтому такое поручение было дано с практически гарантированным исполнением с небольшим вкраплением в виде применения сервисов искусственного интеллекта, опять же без оговорки о сильном его варианте. В любом случае речь идет об именно государственном управлении в сфере оказания электронных услуг с применением сервисов ИИ. Другое дело, насколько в настоящее время возможно придать субъектность человеко-машинным или исключительно машинным решениям, насколько человеческое присутствие в этом будет сокращаться?

В недавнем заключительном докладе Консультативного органа высокого уровня ООН по искусственному интеллекту «Управление искусственным интеллектом в интересах человечества», сформированном в ходе широких консультаций по итогам Саммита будущего (22–23 сентября 2024 г.), были отмечены риски, которые даже при рекомендательном характере документа могут быть учтены в национальном законодательстве стран-участниц.

В докладе подчеркивается, что ИИ преобразует мир и в качестве комплекса технологий обладает огромным потенциалом для благих целей, однако без управления его возможности могут быть не реализованы вовсе или реализованы неравномерно. Среди рисков называется: 1) неспособность делиться передовыми технологиями ИИ ввиду недоверия или неблагоприятных условий; 2) предвзятое использование ИИ не только для слежки, но и в целях конфабуляции больших языковых моделей с помощью ИИ; 3) непрозрачные и автономные системы ИИ бросают вызов традиционным системам регулирования; 4) дефицит управления на глобальном уровне в отношении ИИ, где практика не соответствует риторике, соблюдение норм происходит на добровольной основе, отсутствует подотчетность, в особенности в закрытых системах ИИ, а сама система норм

на стадии зарождения полна пробелов, несмотря на активное обсуждение стандартов и принципов в этой сфере. В докладе называются и другие риски в сфере глобального управления, ускоренного развития ИИ, непонимания всех внутренних механизмов ИИ и ввиду этого его непрогнозируемости и неконтролируемости. Особенно отмечено, что развитие ИИ не может быть отдано только на откуп рынкам [10, с. 7-8].

В исследованиях зарубежных авторов немалое внимание уделяется проблеме эффективной работы «машино-человеческих» команд, под которыми понимают работающие группы от двух и более лиц во взаимозависимых отношениях с одной или несколькими машинами для достижения общей цели. При этом не исключается, что машины иногда дают указания, а иногда принимаются совместные человеко-машинные решения [11, р. 4]. С точки зрения российской практики, такое сочетание, например, на уровне следственных (ст. 163 УПК РФ) и следственно-оперативных групп и даже оперативно-следственных бригад (в последнем случае, согласно классике криминалистики, мы говорим о примате ОРД) такое управление СИИ, как и совместное принятие решений, можно назвать фантастическим. Но соответствующие изменения в законодательстве или, по меньшей мере, в закрытых ведомственных инструкциях в рамках сугубо оперативной работы или деятельности оперативно-следственных групп понадобятся по мере развития СИИ в отечественном сегменте до уровня эффективного управления процессами. Дело в том, что установленные уголовно-процессуальным законодательством общие или частные ограничения к использованию определенных средств и методов неприменимы к СИИ в силу его претензий в зарубежной практике на чуть ли не самостоятельную и даже безальтернативную для познающего субъектность. СИИ, в принципе, априори вне системы правоприменения, поскольку может потенциально

рассматриваться в качестве самостоятельного актора, что вызывает у любого профессионального юриста когнитивный диссонанс. Не за горами время, когда, например, в протоколах постановки диагноза пациентам СИИ может занять едва ли не решающее место. В этой связи возникает три риторических вопроса: 1) какова адаптируемость российской правовой системы под появление в ней нового решающего субъекта всех без исключения отраслей права – СИИ, даже в удобоваримых для нас формах контроля за ним со стороны человека и легитимации окончательным человеческим решением? 2) вытекающий из первого вопрос о том, насколько СИИ в принципе может быть юридически контролируемым? 3) готова ли правовая система к платформизации сервисов правоприменения на симбиозе СИИ, действующего законодательства и практики его правоприменения? В последнем случае прогноз об отмирании права в марксистско-ленинской концепции, пусть и без перехода ко второй фазе коммунизма, может заиграть новыми красками.

Высказываются небезосновательные предложения о том, что необходимо разрабатывать доступный пониманию искусственный интеллект, который будет действовать как можно более гуманно для купирования проблемы человеческого непонимания [11, р. 5]. В то время как доступный пониманию ИИ мог бы помочь в самоуправляемом автомобиле, ситуативно информируя пользователя объяснением того, почему принято то или иное решение. Такого рода сервисы представляются целесообразными при решении посильных человеку аналитических задач, в особенности тогда, когда человек опирается на исключительно машинные предложения. Такое осознание указания к действию со стороны машины выводит рациональность из слепой зоны, но удлиняет процедуру. Однако технологии нередко целенаправленно создаются для того, чтобы «не действовать по-человечески»

для большей эффективности. В разработке ИИ максимальная человекоподобность сугубо технологически не всегда выступает сильной стороной. Биологические процессы торможения на фоне усталости незнакомы машине. Но отрыв от «человеческой матрицы» создает проблемы с доверием и удержанием контроля.

В нормативном регулировании ИИ мы не особо отстаем от США, обращая внимание как на риски, так и на прикладное использование технологии и дополнительные ввиду этого возможности. В 2017 г. искусственный интеллект впервые включен в «Стратегию национальной безопасности США», а годом позже – в Стратегию национальной обороны. В Стратегии национальной безопасности Российской Федерации, утвержденной Указом Президента РФ от 2 июля 2021 г. № 400, искусственный интеллект упоминается трижды: 1) в качестве одной из задач, решение которой среди прочих предполагает достижение цели обеспечения информационной безопасности, заключающейся в совершенствовании соответствующих средств и методов на основе применения передовых технологий, включая технологии искусственного интеллекта и квантовые вычисления (пп. 12. п. 57 Стратегии национальной безопасности РФ; далее – СНБ РФ); 2) в виде задачи повышения производительности труда путем цифровизации, использования технологий искусственного интеллекта, решение которой, наряду с другими, ведет к обеспечению экономической безопасности Российской Федерации (пп. 6 п. 67 СНБ РФ); 3) в качестве задачи развития одной из перспективных высоких технологий – искусственного интеллекта, обработки больших данных, решение которой послужит, по логике составителей, одной из составляющих достижения цели научно-технологического развития Российской Федерации (пп. 14 п. 76 СНБ РФ). Поэтому в последнем случае развитие для развития вызывает вопросы не только по форме, но и по сути:

само по себе «научно-технологическое развитие Российской Федерации» в качестве цели выглядит странно, поскольку нет никаких измеримых параметров, т. е. целью является сам процесс, а не его динамика и определенные контрольные точки, позиции в мире и пр. Например, в Стратегии национальной обороны США ИИ называется в качестве одной из технологий, которые изменяют характер войны и дают все более изощренные возможности противникам, включая негосударственные субъекты [12, с. 325]. И это в 2018 г.! В СНБ РФ в этой части не делается оговорок по поводу СИИ, а сама перспективная и передовая технология находится в подсистеме других прочих даже на уровне решаемых задач.

Одним из максимальных рисков расширения сфер применения СИИ остается потеря творческого потенциала в литературе и даже искусстве. Переход от больших языковых систем к когнитивным моделям как универсальному языку СИИ не исключает гегемонии универсализма при размене на это преимуществ русского языка. Переводы на русский язык английской классики лучше оригинала, и это не фигура речи: К. И. Чуковский переводил произведения О. Генри, Эрнеста Сетона-Томпсона, Марка Твена и др. с «улучшением» как решение осознанной задачи, Б. Пастернак делал превосходящие оригинал переводы английских романтиков Перси Биши Шелли, Джона Китса и Джорджа Гордона Байрона ввиду очевидности большей выразительности и художественной образности русского варианта. Вместе с тем, СИИ-комбинаторика литературного классического наследия «не создаст» нового гениального писателя, когда в принципе по авторству мы получаем нечеловекоцентричное произведение, которое может быть «улучшено» по сюжету, изобразительным средствам, заложенным смыслам. Но это не будет новый гений уровня Ф. М. Достоевского. Можно даже компенсировать отсутствие пейзажистики Ф. М.

Достоевского за счет комбинаторики с творчеством Г. Г. Маркеса, но это не «новый гений», а синтетический кентавр из двух великих писателей. Человеческий производный примат здесь присутствует, но подобный вызов может оказаться непреодолимым для писательской гильдии при либеральной рыночной модели экономики. Творческое вырождение естественного оригинального – вполне прогнозируемая перспектива в меркантилизованной системе ценностей. Сценарии будет дешевле сгенерировать, чем написать в муках человеческого творчества. Выход СИИ за пределы человеческого в литературе и поэзии невозможен созданием подлинно нового. Литературные кентавры и полисемантические по графам произведения не создают подлинно нового гения. Законодательство должно ответить на вызов: «Есть ли место человеческому не СИИ-творчеству и принятию решений, своего рода природному интеллектуальному «хенд-мейд» или «брейн-мейд»?»

Предложения:

1. Закрепление конституционного права не просто на жизнь (ст. 20 Конституции РФ), а на непоколебимый императив легального притязания на творческую жизнь как исключительную прерогативу человека в свободном социокультурном и научном его месторазвитии.

2. Разработка законодательства не только ограничивающего, но и обеспечивающего исследования в сфере сильного ИИ в части обеспечения национальной безопасности (действующая несовершенная Стратегия может быть сохранена как «белый шум») в закрытом варианте НПА, в том числе:

- направление «интеллектуальных экзоскелетов» физической коллаборации человек-машина, т. е. подготовка высококвалифицированных пользователей и аналитиков технологических разработок на основе сильного искусственного интеллекта;

- гипотетически возможное направление генной модификации человека для обретения им свойств сверхразума на основе подобных имеющихся и перспективных биоинженерных технологий;

- комбинаторика двух выше названных форматов.

3. Разработка законодательства в части обеспечения «субъектности» или, по меньшей мере, легитимности для машинных и человеко-машинных решений в риск-ориентированной парадигме. Распределение ответственности «заказчик НИОКР – разработчик СИИ – оператор (пользователь) сервиса СИИ – товары, услуги, госфункции» на риск-ориентированной платформе с учетом приоритетов СЭР.

4. Апробация пилотного проекта по формированию платформенного СИИ-сервиса правореализации в виде готового набора завершено цикла правоприменения, к примеру, в формате рассмотрения споров в сфере закупок.

5. Гармонизация «СИИ-творчества» с авторским правом, главным образом, посредством создания реестра открытий и изобретений (Е. В. Гарин) как сферы защиты человеческого капитала, в том числе с закреплением конституционного права человека на первоисточник.

Библиография

1. Национальная стратегия развития искусственного интеллекта на период до 2030 года, утв. Указом Президента РФ от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» // СПС «КонсультантПлюс».

2. Путин: лидер в сфере искусственного интеллекта станет властелином мира // РИА Новости. 01.09.2017, обновлено 03.03.2020. URL: <https://ria.ru/20170901/1501566046.html> (дата обращения: 19.03.2024).

3. Меркулов И. П. Когнитивные способности. М.: ЦОП Института философии РАН, 2005.

4. ГА ООН приняла первую резолюцию об искусственном интеллекте // Российская газета. 22.03.2024. URL: <https://rg.ru/2024/03/22/ga-oon-priniala-pervuiu-rezoliuciu-ob-iskusstvennom-intellekte.html> (дата обращения: 22.03.2024).

5. Madiega T. Artificial intelligence act. Briefing EU Legislation in Progress. EPRS | European Parliamentary Research Service. March 2024. // URL: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf) (дата обращения: 19.03.2024).

6. ИИ разложили по юридическим полочкам. ЕС согласовал первый в мире закон о регулировании искусственного интеллекта // Коммерсантъ. 10.12.2023. URL: <https://www.kommersant.ru/doc/6394442> (дата обращения: 19.03.2024).

7. Декларация об ответственной разработке и использовании сервисов в сфере генеративного искусственного интеллекта. URL: https://ethics.a-ai.ru/assets/ethics_files/2024/03/13/GenAi_Declaration_Ai_Alliance_Russia_FpNJ2Lc.pdf (дата обращения: 20.03.2024).

8. Кодекс этики в сфере искусственного интеллекта // СПС «КонсультантПлюс».

9. Ведута Н. И. Экономическая кибернетика. Очерки по вопросам теории. Минск: Изд-во «Наука и техника», 1971.

10. Управление искусственным интеллектом в интересах человечества. Advanced Translation Copy. URL: file:///C:/Users/u016/Downloads/governing_ai_for_humanity_final_report_ru.pdf (дата обращения: 24.09.2024).

11. Dubrow Samantha, Orvis Kara L. Human-Machine Teaming: What Skills do the Humans Need? // Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC). 2019. Paper № 20371.

12. Бирюков П. Н. Деятельность США в сфере использования искусственного интеллекта // Международное и европейское право. 2019. № 3.

References

1. National Strategy for the Development of Artificial Intelligence for the period up to 2030, approved by By Decree of the President of the Russian Federation No. 490 dated October 10, 2019 "On the development of artificial intelligence in the Russian Federation". *ConsultantPlus: legal reference system*. (In Russ.).

2. Putin: the leader in the field of artificial intelligence will become the ruler of the world. *RIA Novosti*. 01.09.2017, upd. 03.03.2020. Available at: <https://ria.ru/20170901/1501566046.html> (accessed: 19.03.2024). (In Russ.).

3. Merkulov I. P. *Cognitive abilities*. Moscow: TSOP Institute of Philosophy of the Russian Academy of Sciences Publ.; 2005. (In Russ.).
4. The UN General Assembly adopted the first resolution on artificial intelligence. *Rossiyskaya Gazeta*. 22.03.2024. Available at: <https://rg.ru/2024/03/22/ga-oon-priniala-pervuiu-rezoliuciiu-ob-iskusstvennom-intellekte.html> (accessed: 22.03.2024) (In Russ.).
5. Madiaga T. Artificial intelligence act. Briefing EU Legislation in Progress. *EPRS | European Parliamentary Research Service*. March 2024. Available at: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf) (accessed: 19.03.2024).
6. AI was laid out on legal shelves. The EU has agreed on the world's first law on the regulation of artificial intelligence. *Kommersant*. 10.12.2023. Available at: <https://www.kommersant.ru/doc/6394442> (accessed: 19.03.2024). (In Russ.).
7. Declaration on responsible development and use of services in the field of generative artificial intelligence. Available at: https://ethics.a-ai.ru/assets/ethics_files/2024/03/13/GenAi_Declaration_Ai_Alliance_Russia_FpNJ2Lc.pdf (accessed: 20.03.2024). (In Russ.).
8. Code of ethics in the field of artificial intelligence. *ConsultantPlus: legal reference system*. (In Russ.).
9. Veduta N. I. *Economic cybernetics. Essays on the theory*. Minsk: Nauka i tehnika Publ.; 1971. (In Russ.).
10. Artificial intelligence management in the interests of humanity. Advanced Translation Copy. Available at: file:///C:/Users/u016/Downloads/governing_ai_for_humanity_final_report_en.pdf (accessed: 24.09.2024).
11. Dubrow Samantha, Orvis Kara L. Human-Machine Teaming: What Skills do the Humans Need? *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*. 2019; Paper No. 20371.
12. Biryukov P. N. U.S. activities in the field of artificial intelligence. *Mezhdunarodnoe i evropejskoe pravo*. 2019; 3. (In Russ.).

Информация об авторах

Полстовалов Олег Владимирович, доктор юридических наук, профессор, эксперт Сretenского клуба, АНО «Содействие и развитие инноваций в научно-производственной сфере», г. Москва, Россия, e-mail: Polstovalov74@mail.ru

Information about the authors

Oleg V. Polstovalov, Doctor of Law, Professor, Sretensky Club expert, ANO «Assistance and Development of Innovations in the Scientific and Industrial Sphere», Moscow, Russia, e-mail: Polstovalov74@mail.ru

Для цитирования

Полстовалов О. В. Перспективы правового регулирования в сфере сильного искусственного интеллекта и государственное управление: попытки ограничения безграничного // Журнал Высокотехнологичное право. – 2026. Т. 2, № 1 (3). – С. 95-119

For citation

Polstovalov O. V. Prospects for legal regulation in the field of strong artificial intelligence and public administration: attempts to limit the limitless // Journal of High-tech Law. – 2026. Vol. 2, No. 1 (3). – Pp. 95-119

Поступила в редакцию / Received 02.03.2026

Поступила после рецензирования / Received after review 26.03.2026

Принята к публикации / Accepted 27.03.2026