П О С Т А Н О В Л Е Н И Е

Парламентской Ассамблеи
Организации Договора о коллективной безопасности

О Рекомендациях для государств – членов ОДКБ по выработке общих принципов развития национального законодательства в области создания искусственного интеллекта и робототехники в целях обеспечения национальной безопасности

Парламентская Ассамблея Организации Договора о коллективной безопасности п о с т а н о в л я е т:

1. Принять Рекомендации для государств – членов ОДКБ по выработке общих принципов развития национального законодательства в области создания искусственного интеллекта и робототехники в целях обеспечения национальной безопасности (далее – Рекомендации) (прилагаются).

2. Направить Рекомендации в парламенты государств – членов ОДКБ для использования в работе по совершенствованию законодательства государств – членов Организации в соответствующей сфере.

3. Опубликовать текст Рекомендаций на официальном сайте и в материалах Парламентской Ассамблеи ОДКБ.

Председатель
Парламентской Ассамблеи ОДКБ **В.В.ВОЛОДИН**

Москва
9 декабря 2024 года
№ 17-7.5

Приложение
к постановлению ПА ОДКБ
от 9 декабря 2024 года № 17-7.5

РЕКОМЕНДАЦИИ
для государств – членов ОДКБ по выработке общих принципов развития национального законодательства в области создания искусственного интеллекта и робототехники в целях обеспечения национальной безопасности

В настоящее время активно развиваются технологии искусственного интеллекта и робототехники (далее – ИИ). Особенно заметен прогресс в области так называемого генеративного ИИ (GPT), достигнутый в 2022–2023 годах и связанный с появлением систем 3-го и 4-го поколений, а также предоставлением возможности доступа к данным технологиям широкому кругу пользователей.

Национальные программы развития ИИ приняты более чем в 60 странах, в ряде стран ведется разработка подобных стратегий.

Имея очевидные преимущества для развития экономики, государственного управления и оборонной сферы, технологии ИИ создают целый ряд угроз безопасности граждан, общества и государства. Среди таких угроз обычно отмечается использование ИИ для осуществления сетевых атак, создания фейковой и иной недостоверной (фальсифицированной) информации, имитации человека. Другая группа угроз связана с недостаточной прозрачностью, понятностью и предсказуемостью систем ИИ. Применение технологий ИИ для принятия управленческих решений создает угрозу дискриминации граждан по признаку расы, пола и др. Одной из активно развивающихся и одновременно вызывающих особую озабоченность технологий является технология биометрической идентификации, позволяющая идентифицировать человека по фото- или видеоизображению, а также по голосу.

В большинстве постсоветских стран развитие ИИ не выделяется в самостоятельное направление государственной политики, однако упоминания об этих технологиях присутствуют в стратегических документах с более широкой сферой (концепции научно-технического развития, цифрового государства и т. п.).

Вместе с тем выработка единой политики в области ИИ особенно важна для проведения согласованной политики в сфере обороны и безопасности в рамках ОДКБ.

Целью настоящих Рекомендаций является формирование обоснованных предложений по выработке общих принципов (направлений) развития национального законодательства государств – членов ОДКБ в области использования ИИ для обеспечения национальной безопасности.

Степень разработанности нормативного регулирования ИИ в государствах – членах ОДКБ

Республика Армения

В информационной сфере наиболее значимыми являются Закон Республики Армения от 22 октября 2003 года № ЗР-11 «О свободе информации», Закон Республики Армения от 30 декабря 2014 года № ЗР-245 «О государственном содействии в сфере информационных технологий», Закон Республики Армения от 13 июня 2015 года № ЗР-49 «О защите персональных данных», Программа Правительства Республики Армения (приложение к Постановлению Правительства Республики Армения № 1363-А от 18 августа 2021 года), в разделе 1.3 которой, посвященном сфере обороны, ставятся задачи по «внедрению единой автоматизированной системы управления… технологической модернизации Вооруженных Сил и применению наделенных искусственным интеллектом систем», а в разделе 4.4, посвященном науке, ИИ отнесен к приоритетным направлениям и предусмотрено дополнительное финансирование проектов в данной сфере.

Республика Беларусь

Основным законом, регулирующим отношения в информационной сфере, является Закон Республики Беларусь от 10 ноября 2008 года № 455-З «Об информации, информатизации и защите информации». Однако он не содержит положений, непосредственно посвященных системам ИИ.

Декрет Президента Республики Беларусь от 22 сентября 2005 года № 12 «О Парке высоких технологий» (редакция 2023 года) предусматривает особый правовой режим («регуляторную песочницу») для резидентов Парка – организаций, ведущих деятельность в сфере ИКТ, в том числе наделяет их правом вести деятельность «в сфере искусственного интеллекта, создания систем беспилотного управления транспортными средствами».

Среди документов стратегического развития, затрагивающих сферу ИИ, следует назвать Концепцию информационной безопасности Республики Беларусь, утвержденную Постановлением Совета Безопасности Республики Беларусь от 18 марта 2019 года № 1; Государственную программу «Цифровое развитие Беларуси» на 2021–2025 годы, утвержденную Постановлением Совета Министров Республики Беларусь от 2 февраля 2021 года № 66; Государственную программу инновационного развития Республики Беларусь на 2021–2025 годы, утвержденную Указом Президента Республики Беларусь от 15 сентября 2021 года № 348.

Определение термина «искусственный интеллект» содержится в Постановлении Совета Министров Республики Беларусь «О мерах по реализации Указа Президента Республики Беларусь от 7 апреля 2022 года № 136»: «Искусственный интеллект – комплекс технологических решений, позволяющий имитировать когнитивные функции человека (в том числе самообучение и поиск решений без заранее заданного алгоритма) и получать

при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека, и включающий в себя информационно-коммуникационную инфраструктуру, программное обеспечение, процессы и сервисы по обработке данных и поиску решений».

Республика Казахстан

Закон Республики Казахстан от 24 ноября 2015 года № 418-V ЗРК «Об информатизации» является основным документом, регулирующим отношения в информационной сфере. В законе не дается понятие ИИ, однако раскрываются такие понятия, как:

национальная платформа искусственного интеллекта – технологическая платформа, предназначенная для сбора, обработки, хранения и распространения наборов данных и предоставления услуг в области искусственного интеллекта;

оператор национальной платформы искусственного интеллекта – юридическое лицо, определяемое Правительством Республики Казахстан, на которое возложено обеспечение развития и функционирования закрепленной за ним Национальной платформы искусственного интеллекта;

интеллектуальный робот – автоматизированное устройство, совершающее определенное действие или бездействующее с учетом воспринятой и распознанной внешней среды (пункты 54, 55 и 43-1 статьи 1 соответственно).

Также указанный закон определяет ряд более общих категорий, важных с точки зрения регулирования технологий ИИ: информационная система, объект информатизации и др. Статья 13-2 закрепляет правовой статус оператора национальной платформы ИИ; статья 18-1 – права и обязанности собственника и владельца интеллектуального робота.

Закон Республики Казахстан от 10 июля 2023 года № 18-VIII ЗРК «Об онлайн-платформах и онлайн-рекламе» закрепил обязанность собственника онлайн-платформы «принимать меры по совершенствованию... алгоритмов искусственного интеллекта» (статья 13).

В Республике Казахстан развитие технологий ИИ и анализа больших данных выделено в качестве одного из основных приоритетов, особенно в областях экономики, обеспечения безопасности, медицины, политики (Постановление Правительства Республики Казахстан от 28 марта 2023 года № 269 «Об утверждении Концепции цифровой трансформации, развития отрасли информационно-коммуникационных технологий и кибербезопасности на 2023–2029 годы»).

В Концепции цифровой трансформации, развития отрасли информационно-коммуникационных технологий и кибербезопасности на 2023–2029 годы непосредственно вопросам ИИ не уделяется большого внимания, однако на 2023 год запланирована разработка дорожной карты (стратегическое видение) по развитию искусственного интеллекта, на 2024 год

– создание национальной системы искусственного интеллекта на базе Smart Data Ukimet, которая позволит прогнозировать и принимать решения на основе достоверных данных. Тем не менее Концепция предусматривает целый комплекс мероприятий в сфере информатизации, которые неизбежно окажут существенное влияние и на сегмент, связанный с технологиями ИИ. Среди таких мероприятий – агрегация актуальных данных, собираемых на основе Единой платформы «электронного правительства».

Концепция развития искусственного интеллекта на 2024–2029 годы утверждена Постановлением Правительства Республики Казахстан от 24 июля 2024 года № 592. В Концепции сформулированы принципы и подходы к развитию ИИ, в том числе серьезное внимание уделяется вопросам правового регулирования отношений в сфере ИИ. Концепция предусматривает разработку нормативных правовых актов, в которых будут отражены понятийный аппарат, основные этические принципы применения ИИ, нормы по безопасности применения ИИ и другие вопросы (глава 2 раздела 5).

В Республике Казахстан ведется разработка проекта Цифрового кодекса. В ноябре 2023 года был опубликован Консультативный документ регуляторной политики к проекту Цифрового кодекса Республики Казахстан, в апреле 2024 года – текст проекта (https://online.zakon.kz/Document/?doc_id=38933548).

Кыргызская Республика

В числе наиболее значимых актов, регулирующих отношения в информационной сфере, следует назвать Закон Кыргызской Республики от 19 июля 2017 года № 127 «Об электронном управлении» и Указ Президента Кыргызской Республики от 17 декабря 2020 года № 64 «О неотложных мерах по активизации внедрения цифровых технологий в государственное управление Кыргызской Республики».

Решением Совета безопасности Кыргызской Республики от 14 декабря 2018 года № 2 одобрена Концепция цифровой трансформации «Цифровой Кыргызстан 2019–2023», распоряжением Кабинета Министров Кыргызской Республики от 12 января 2022 года № 2-р утвержден План мероприятий по цифровизации управления и развития цифровой инфраструктуры в Кыргызской Республике на 2022–2023 годы. Специальное регулирование отношений в сфере ИИ в настоящее время отсутствует.

Важнейшим событием в рассматриваемой сфере стало опубликование в августе 2023 года проекта Цифрового кодекса Кыргызской Республики (http://koomtalkuu.gov.kg/ru/view-npa/2927). Проект содержит отдельную главу, посвященную регулированию систем ИИ. В ней закрепляются принципы регулирования ИИ, устанавливается обязательная оценка опасности, которую несет охраняемым благам создаваемая система ИИ. Системы ИИ с повышенной опасностью (определенные как системы, использование которых повышает вероятность причинения вреда охраняемым благам) выделяются в отдельную категорию, в отношении которой действуют

специальные правила, относящиеся к управлению рисками, обеспечению открытости и объяснимости систем ИИ повышенной опасности, их подконтрольности, точности, надежности и цифровой устойчивости (статьи 190, 191). Закрепляются обязанности пользователей систем ИИ повышенной опасности (статья 192).

Российская Федерация

Из числа государств, входящих в ОДКБ, стратегические акты для технологий ИИ приняты только в России: Национальная стратегия развития ИИ (утверждена Указом Президента Российской Федерации от 10 октября 2019 года № 490, действует в редакции Указа Президента Российской Федерации от 15 февраля 2024 года № 124) и Концепция развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 года (утверждена распоряжением Правительства Российской Федерации от 19 августа 2020 года № 2129-р).

Национальная стратегия развития искусственного интеллекта на период до 2030 года (далее – Стратегия) определяет основные принципы развития и использования технологий ИИ, в том числе принципы защиты прав и свобод человека; безопасности; прозрачности; технологического суверенитета; целостности инновационного цикла; наиболее эффективного использования технологий ИИ; поддержки конкуренции; открытости и доступности; преемственности; защищенности; достоверности исходных данных.

В Стратегии целями развития ИИ в Российской Федерации названы: 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 в области ИИ.

Отдельным направлением Стратегии является создание комплексной системы регулирования общественных отношений, возникающих в связи с развитием и использованием технологий ИИ. Основной целью данного направления Стратегии (в редакции от 15 февраля 2024 года) названо создание в Российской Федерации благоприятных нормативно-правовых условий для разработки, внедрения и использования технологий ИИ и решений, разработанных на их основе, с учетом обеспечения защиты прав и свобод человека и безопасности Российской Федерации. Сформулированы принципы нормативно-правового регулирования общественных отношений, связанных с развитием и использованием технологий ИИ: безопасность, гуманистический подход, уважение автономии и свободы воли человека, недискриминация, риск-ориентированный подход, ответственность, квалифицированная экспертная оценка.

В развитие Стратегии разработан и принят в 2020 году федеральный проект «Искусственный интеллект», содержащий дорожную карту конкретных мероприятий и плановые ключевые показатели до 2024 года. Президент Российской Федерации В.В.Путин поручил включить федеральный

проект «Искусственный интеллект» в национальный проект по формированию экономики данных. Таким образом, он будет продлен до 2030 года.

Концепция развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 года (далее – Концепция) стала основополагающим документом, устанавливающим принципы и направления развития регулирования ИИ и робототехники (РТ) в России. В соответствии с опубликованным текстом Концепции приоритетной целью регулирования отношений в сфере ИИ и РТ на данном этапе их развития является стимулирование разработки, внедрения и использования безопасных и заслуживающих доверия технологий и систем ИИ и РТ для их применения во благо человека, общества и государства. Регулирование в сфере ИИ должно способствовать достижению высоких темпов экономического роста, повышению благосостояния и качества жизни граждан, обеспечению национальной безопасности и правопорядка, достижению устойчивой конкурентоспособности российской экономики, в том числе лидирующих позиций в мире в области ИИ.

В качестве основных общеотраслевых задач регулирования применения технологий ИИ и РТ в Концепции указаны создание механизмов упрощенного внедрения продуктов с использованием технологий ИИ и РТ, установление юридической ответственности в случае применения систем ИИ и РТ, развитие страховых институтов, совершенствование режима оборота данных, режима экспорта систем ИИ и РТ, системы технического регулирования и оценки соответствия.

Концепцией определены отраслевые направления совершенствования регулирования применения технологий ИИ и РТ, среди которых: сфера охраны здоровья граждан, государственное (муниципальное) управление, транспорт, градостроительство, концепция «умного города», финансовая сфера, космическая деятельность, промышленность.

Документ также предусматривает регуляторные меры для финансового стимулирования развития отрасли. Необходимо оценить целесообразность и по итогам такой оценки проработать меры поддержки по четырем направлениям: стимулирование предложения, стимулирование спроса, развитие государственно-частного партнерства, развитие экспорта.

Важно подчеркнуть, что документ призван поставить в центр внимания регулятора человека: его права и интересы, его безопасность и благополучие в контексте внедрения новых технологий. В отличие от ряда зарубежных инициатив, в документе даже не ставится вопрос о наделении какими-либо правами роботов или юнитов ИИ. Развитие технологий всегда должно быть во благо граждан и оставаться под полным контролем человека.

В Российской Федерации действует законодательство об экспериментальных правовых режимах, которые являются важным механизмом тестирования технологий ИИ. Приняты федеральные законы от 24 апреля 2020 года № 123-ФЗ «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для

разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона "О персональных данных"» и от 31 июля 2020 года № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации» (с 1 января 2025 года вступает в силу новая редакция с изменениями, внесенными Федеральным законом от 8 июля 2024 года № 169-ФЗ, которая содержит ряд положений, касающихся случаев причинения вреда жизни, здоровью или имуществу других лиц при реализации экспериментального правового режима, в том числе в результате использования решений, разработанных с применением технологий искусственного интеллекта).

Также в Российской Федерации утвержден комплекс национальных стандартов в области ИИ, который включает 111 документов по направлениям «Межотраслевые стандарты», «Общие вопросы качества», «Транспорт», «Здравоохранение», «Образование», «Данные», «Сельское хозяйство», «Промышленность» (по состоянию на 29 августа 2024 года).

29 декабря 2023 года Минэкономразвития России и Росстандартом утверждена Перспективная программа стандартизации по приоритетному направлению «Искусственный интеллект» на 2021–2024 годы в обновленной редакции. Предусмотрена разработка 131 технического стандарта по направлениям «Межотраслевые стандарты», «Данные», «Информационная инфраструктура», «Качество», «Варианты использования», «Автомобильный транспорт», «Водный транспорт», «Образование», «Здравоохранение», «Промышленность», «Средства измерения», «Строительство», «Пространственные данные, дистанционное зондирование Земли», «Специализированная техника».

Правоприменительная практика, связанная с применением технологий ИИ, на сегодняшний день в Российской Федерации состоит в основном из дел, касающихся вопросов интеллектуальной собственности в случаях, когда объект создан с применением технологии дипфейков, нейросетей, а также касающихся использования систем на основе ИИ для осуществления телефонных звонков либо в рекламных целях, либо в рамках деятельности по возврату просроченной задолженности. Такие действия квалифицируются по статье 14.57 Кодекса об административных правонарушениях Российской Федерации.

Среди актов рекомендательного характера следует указать Кодекс этики в сфере искусственного интеллекта, разработанный в 2021 году, к которому в настоящее время присоединилось более 380 подписантов из 21 юрисдикции (в том числе Беларусь, Куба, Таджикистан, страны Африки и другие).

Республика Таджикистан

К числу основных законов в информационной сфере в Республике Таджикистан следует отнести Закон Республики Таджикистан от 6 августа 2001 года № 40 «Об информатизации» (действует в редакции 2022 года) и

Закон Республики Таджикистан от 3 августа 2018 года № 1537 «О защите персональных данных».

Концепция цифровой экономики в Республике Таджикистан, утвержденная Постановлением Правительства Республики Таджикистан от 30 декабря 2019 года № 642, содержит ряд упоминаний технологий ИИ. Наиболее важным представляется план по разработке стандартов информационной безопасности в системах, реализующих технологии ИИ и обеспечение надзора за их соблюдением. Также упоминаются создание интеллектуальной энергосистемы, интеллектуальной транспортной системы, внедрение ИИ в медицине, городском хозяйстве (Smart City).

Мировой опыт: регулирование ИИ на международном уровне и в других государствах

Организация экономического сотрудничества и развития

Важнейшую роль в формировании подходов к регулированию ИИ сыграла Организация экономического сотрудничества и развития (ОЭСР), подготовившая в 2019 году Рекомендации по искусственному интеллекту[1]. В рекомендациях получили закрепление пять принципов управления ИИ:

1) инклюзивный рост, устойчивое развитие и благосостояние;
2) ориентированность на человеческие ценности и справедливость;
3) прозрачность и объяснимость;
4) надежность и безопасность;
5) контролируемость.

ЮНЕСКО

Рекомендация ЮНЕСКО об этических аспектах искусственного интеллекта от 23 ноября 2021 года[2] является в настоящее время одним из немногих рекомендательных актов, касающихся гуманитарных проблем внедрения технологий ИИ. В качестве этических и иных принципов, на которых должны выстраиваться системы ИИ, в данном акте выделяются:

1) соразмерность и непричинение вреда;
2) безопасность и защищенность;
3) справедливость и отказ от дискриминации;
4) устойчивость;
5) право на неприкосновенность частной жизни и защита данных;
6) подконтрольность и подчиненность человеку;
7) прозрачность и объяснимость;
8) ответственность и подотчетность;

---

[1] Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449. URL: https://legalinstruments.oecd.org/api/print?ids=648&lang=en.

[2] https://unesdoc.unesco.org/ark:/48223/pf0000380455_rus.locale=ru

9) многостороннее и адаптивное управление и взаимодействие.

Принцип безопасности и защищенности сформулирован как необходимость «учитывать, предотвращать и ликвидировать» два вида рисков: риски для безопасности и риски для защищенности. Первые связаны с непреднамеренным причинением вреда, вторые – с уязвимостью для кибератак. Отмечается, что безопасность и защищенность ИИ-систем могут быть повышены посредством разработки надежных и защищенных от несанкционированного доступа к личной информации комплексных систем, которые обеспечат более эффективный характер обучения и сертификации моделей ИИ на основе качественных данных.

Под прозрачностью понимается право получать информацию о том, какие решения принимаются с использованием алгоритмов ИИ или на основе полученных с их помощью данных, в том числе решения, затрагивающие безопасность и права человека, чтобы в соответствующих случаях была возможность запросить у субъектов ИИ или у государственных учреждений пояснительную информацию. Объяснимость касается обеспечения понятности и предоставления разъяснений в отношении полученных с помощью ИИ-систем результатов, а также доступности для понимания исходных данных, непосредственных результатов, функционирования каждой алгоритмической структуры и того, как все это влияет на выданный системой результат.

В Рекомендации ЮНЕСКО названы 11 приоритетных областей, которые, по мнению разработчиков, требуют принятия стратегических мер. Это: оценка этического воздействия; этичное управление и руководство; политика в отношении данных; развитие и международное сотрудничество; окружающая среда и экосистемы; гендерное равенство; культура; образование и научные исследования; коммуникация и информация; экономика и рынок труда; здоровье и социальное благополучие.

Европейский союз

1 августа 2024 года вступил в силу Регламент Европейского союза об искусственном интеллекте (Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)).

Регламент предусматривает разделение систем ИИ на четыре категории в зависимости от уровня риска: с ограниченным, высоким и неприемлемым риском, а также генерирующие системы (нейросети). Основное внимание уделяется системам с высоким риском.

Статья 5 Регламента содержит перечень запретов в отношении систем ИИ. Так, запрещено размещение на рынке и использование систем ИИ, которые: воздействуют на подсознание или манипулируют поведением людей;

используют уязвимости человека с целью существенного искажения его поведения, если это причинило или могло причинить ему существенный вред.

Регламент предусматривает запрет на создание и эксплуатацию систем удаленной биометрической идентификации в общедоступных местах в режиме реального времени в целях обеспечения соблюдения закона, делая исключение для применения этой технологии в целях расследования таких преступлений, как похищение людей или торговля людьми, поиска пропавших граждан, предотвращения терактов и иных угроз жизни и физической безопасности. Регламент также устанавливает ряд запретов в отношении систем ИИ, связанных с технологиями распознавания лиц из Интернета или видеозаписи с камер видеонаблюдения, технологиями определения эмоций физического лица и др.

Главным критерием отнесения системы ИИ к высокорисковым является наличие «гармонизирующего законодательства Европейского союза» в отношении продукта, в котором используется система ИИ, и одновременно наличие требований об обязательной оценке продукта третьей стороной. Кроме того, Регламент содержит Приложение III, в котором прямо перечислены некоторые категории высокорисковых систем: биометрические системы; системы, используемые как компонент безопасности на транспорте, в электро-, водо-, газоснабжении; в сфере образования – для целей приема или оценки учащихся; в сфере управления персоналом; для оказания государственных услуг; системы, используемые для оценки кредитоспособности; системы, используемые правоохранительными органами (детекторы лжи, системы проверки доказательств; системы для анализа преступлений), органами пограничного контроля.

Особо следует отметить системы ИИ, предназначенные для оказания влияния на исход выборов или референдума либо на поведение физических лиц при голосовании на выборах или референдумах. Также к высокорисковым отнесены системы ИИ, предназначенные для использования платформами крупных социальных сетей.

Для высокорисковых систем Регламент закрепляет обязанность разработать, внедрить, документировать и поддерживать систему управления рисками; принимать меры, направленные на устранение (если это возможно) рисков или на снижение их уровня. Системы ИИ с высоким уровнем риска должны быть протестированы для выявления наиболее подходящих и целенаправленных мер по управлению рисками и сопоставления любых таких мер с потенциальными преимуществами и предполагаемыми целями системы. Тестирование должно гарантировать, что системы ИИ с высоким уровнем риска стабильно работают по назначению и соответствуют установленным требованиям.

Статья 13 Регламента закрепляет требования прозрачности и понятности систем: пользователь должен иметь возможность понимать и использовать систему ИИ надлежащим образом, зная в целом, как работает эта система и какие данные она обрабатывает, что позволяет пользователю объяснять

решения, принятые системой ИИ, пострадавшему лицу. Системы ИИ высокого риска должны сопровождаться понятными инструкциями по использованию в соответствующем цифровом формате или иным образом доступными на постоянном носителе, включающими краткую, правильную, четкую и, насколько это возможно, полную информацию, которая помогает в эксплуатации и обслуживании системы ИИ, а также принятии обоснованных решений пользователями, являясь достаточно актуальной, доступной и понятной для пользователей. В пункте 3 данной статьи закрепляется минимальный перечень сведений, которые должны быть отражены в указанных инструкциях.

Регламент закрепляет обязательное встраивание в систему ИИ инструментов человеческого надзора, включая возможность вмешиваться в работу системы ИИ высокого риска или прерывать ее работу (статья 14).

Высокорисковые системы ИИ подлежат обязательной регистрации (статьи 16, 51, 60 и Приложение VIII). Информация, содержащаяся в базе данных Европейского союза, должна быть доступна для общественности.

Содружество Независимых Государств

14 апреля 2023 года на пятьдесят пятом пленарном заседании Межпарламентской Ассамблеи государств – участников Содружества Независимых Государств были приняты Рекомендации по нормативному регулированию использования искусственного интеллекта, включая этические стандарты для исследований и разработок (постановление № 55-23 от 14 апреля 2023 года). Цель данного документа – определение единых основ формирования системы правовых и этических норм, призванных стимулировать разработку и применение систем ИИ в различных отраслях экономики государств – участников СНГ с соблюдением прав граждан и обеспечением безопасности личности, общества и государства. Подразумевается, что Рекомендации будут играть роль дорожной карты для начала формирования системы модельных законов, а также свода морально-этических норм (модельных кодексов и конвенций). В Рекомендациях определены приоритетные задачи для развития сферы ИИ в государствах СНГ, основные направления регулирования ИИ на пространстве Содружества, а также сформулированы следующие принципы регулирования ИИ: защита прав граждан; минимизация рисков; открытость для технологического развития; риск-ориентированный подход; профессиональная оценка; баланс интересов; технологический суверенитет; исключительность военной сферы; поддержка конкуренции; расширение сферы саморегулирования; гармонизация законодательства.

Китай

Комплексное регулирование отношений в сфере ИИ на общегосударственном уровне в настоящее время в Китае отсутствует. Из наиболее значимых актов, регулирующих информационную сферу, следует

указать законы о защите личной информации 2021 года, о кибербезопасности 2017 года и о безопасности данных 2021 года. Кроме того, на локальном уровне приняты Положение о содействии индустрии искусственного интеллекта в Шэньчжэньской особой экономической зоне и Положение о содействии развитию отрасли искусственного интеллекта в муниципалитете Шанхая[3].

В документах стратегического планирования в Китае вопросам ИИ уделяется существенное внимание начиная с 2015 года. Из наиболее значимых документов можно назвать такие акты, которые в государствах – членах ОДКБ принято относить к категории документов стратегического планирования:

– «Сделано в Китае 2025» (май 2015 года);

– «Интернет+» (июль 2015 года);

– Трехлетний план действий в области ИИ (май 2016 года);

– План развития искусственного интеллекта нового поколения (июль 2017 года);

– Имплементация видения в провинции Шанхай развития искусственного интеллекта нового поколения (ноябрь 2017 года);

– Некоторые меры содействия развитию индустрии ИИ (ноябрь 2017 года);

– Новые руководящие принципы для развития индустрии ИИ (декабрь 2018 года);

– Белая книга по стандартизации ИИ (январь 2018 года).

В Китае 10 июля 2023 года были одобрены и с 15 августа 2023 года вступили в силу Временные меры по управлению услугами генеративного искусственного интеллекта[4]. Генеративную технологию ИИ они определяют как модели и связанные с ними технологии, способные генерировать контент, такой как текст, изображения, аудио и видео.

Статья 4 Временных мер устанавливает, что предоставление продуктов или услуг генеративного ИИ должно соответствовать требованиям законов и нормативных актов, уважать общественную мораль и этику. В этой статье закреплены следующие специальные требования: 1) недопустимость создания контента, который «подстрекает к подрыву государственной власти и свержению социалистической системы, угрожает национальной безопасности и интересам… пропагандирует терроризм, экстремизм, пропагандирует этническую ненависть, этническую дискриминацию, насилие, непристойность, а также ложную и вредную информацию, запрещенную законами и административными постановлениями»; 2) недопустимость дискриминации; 3) уважение прав интеллектуальной собственности, соблюдение принципа добросовестной конкуренции; 4) уважение законных

---

[3] Schildkraut P., Zhang H. What To Know About China's New AI Regulations. Arnold & Porter, 2023. URL: https://www.arnoldporter.com/-/media/files/perspectives/publications/2023/04/what-to-know-about-chinas-new-ai-regulations.pdf.

[4] http://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm

прав и интересов других, безопасность для здоровья, соблюдение прав на портрет, на репутацию, права на честь, права на неприкосновенность частной жизни и права на личную информацию других; 5) обязанность провайдера принимать эффективные меры для повышения прозрачности генерирующих услуг искусственного интеллекта и повышения точности и надежности генерируемого контента.

Кроме того, контент, созданный при помощи генеративных технологий ИИ, должен иметь соответствующую маркировку. Закреплена обязанность провайдера создать механизм приема жалоб от пользователей (статья15), а также установлено право пользователей на подачу жалобы в соответствующие компетентные органы (статья 18).

По сообщениям СМИ, по состоянию на январь 2024 года около 40 моделей генеративного ИИ прошли официальную оценку соответствия согласно Временным мерам.

Бразилия

В Бразилии с 2019 года по настоящее время разработаны четыре варианта проекта закона об ИИ. Последний из них, внесенный в мае 2023 года, заменил три предыдущие версии. Основными в проекте являются принципы добросовестности, самоопределения и свободы выбора; прозрачности, объяснимости, понятности, прослеживаемости и возможности аудита; участия человека и контроля; недискриминации; справедливости и инклюзивности; надежности и устойчивости, информационной безопасности.

Бразильский проект, так же как и проект Регламента Европейского союза, предусматривает категорирование систем ИИ по степени риска, закрепляет права лиц, затронутых системами ИИ (независимо от классификации риска системы ИИ): право на информацию об их взаимодействии с системой ИИ до ее использования; право на объяснение решения, рекомендации или прогноза, данных системой ИИ; право оспаривать решения или прогнозы систем ИИ, которые имеют юридические последствия или существенно влияют на интересы пострадавшей стороны; право на вмешательство человека в решения, принимаемые исключительно системами ИИ; право на недискриминацию и исправление дискриминационных предубеждений; право на неприкосновенность частной жизни и защиту персональных данных.

Понятийный аппарат

Рекомендации ОЭСР по искусственному интеллекту определили систему ИИ как машинную систему, которая может для заданного набора целей, определенных человеком, делать прогнозы, давать рекомендации или принимать решения, влияющие на реальную или виртуальную среду. При этом системы искусственного интеллекта предназначены для работы с различными уровнями автономности.

Рекомендация ЮНЕСКО об этических аспектах искусственного интеллекта от 23 ноября 2021 года определяет ИИ следующим образом:

системы на основе ИИ – технологические системы, способные обрабатывать данные и информацию способом, напоминающим разумное поведение и включающим, как правило, такие аспекты, как рассуждение, обучение, распознание, прогнозирование, планирование и контроль.

Существенное значение в таком подходе имеют следующие характеристики:

1. Системы на основе ИИ представляют собой технологии обработки информации, которые включают модели и алгоритмы, обеспечивающие способность обучения и выполнения когнитивных задач, с получением результатов в виде прогнозной оценки и принятия решения в материальной и виртуальной среде.

2. ИИ-системы предназначены для работы с той или иной долей автономности посредством моделирования и представления знаний, а также использования данных и расчета корреляционных зависимостей. В системах на основе ИИ могут использоваться различные методологии, в частности:

– самообучение машины, включая глубокое обучение и обучение с подкреплением;

– автоматизированное рассуждение, включая планирование, диспетчеризацию, представление знаний и формирование рассуждений, поиск и оптимизацию.

3. Системы на основе ИИ могут использоваться в киберфизических системах, включая системы контроля оборудования через Интернет, робототехническое оборудование, социальную робототехнику и системы человеко-машинного интерфейса, объединяющие в себе функции контроля, распознания, обработки данных, собранных датчиками, а также работу исполнительных элементов в среде функционирования ИИ-систем.

Однако в Рекомендации прямо указывается, что преследовалась цель привлечь внимание к ключевым характеристикам ИИ, а не дать единственно возможное определение.

Проект Регламента Европейского союза в редакции от 21 апреля 2021 года определял систему ИИ как «программное обеспечение, которое разработано с использованием одного или нескольких методов и подходов, перечисленных в Приложении I, и может для заданного набора целей, определенных человеком, генерировать такие выходные данные, как контент, прогнозы, рекомендации или решения, влияющие на среду, с которой они взаимодействуют».

Регламент в утвержденной редакции определил систему ИИ как систему, основанную на использовании технических средств, предназначенную для работы с различными уровнями автономии и способную для явных или неявных целей генерировать результаты, такие как прогнозы, рекомендации или решения, влияющие на физическое состояние, или виртуальные среды.

Предлагаемая дефиниция подробно разъясняется в пункте 12 преамбулы Регламента, где отмечается, что она должна быть четкой и одновременно гибкой, чтобы иметь возможность адаптироваться к новым развивающимся

технологиям. Ключевыми особенностями, позволяющими отличать ИИ от иных программных систем, называются «способность к обучению, рассуждению или моделированию» и определенная степень автономии (способность работать без контроля или вмешательства человека).

Национальная стратегия развития искусственного интеллекта на период до 2030 года Российской Федерации в первоначальной редакции определяла ИИ как «комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе, в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений». Такое же определение содержится в Федеральном законе от 24 апреля 2020 года № 123-ФЗ «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона “О персональных данных”».

Указом Президента Российской Федерации от 15 февраля 2024 года в Национальную стратегию развития искусственного интеллекта были внесены изменения, и новое определение ИИ сформулировано 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. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений»; технологии ИИ определены как «совокупность технологий, включающая в себя компьютерное зрение, обработку естественного языка, распознавание и синтез речи, интеллектуальную поддержку принятия решений и перспективные методы искусственного интеллекта».

Выводы

1. Правовое регулирование ИИ во всем мире находится на начальном этапе. Единственный действующий нормативный акт, системно регулирующий сферу ИИ в настоящее время, – Регламент Европейского союза об искусственном интеллекте. В ряде стран действуют отдельные нормы, посвященные ИИ. Среди стран – членов ОДКБ в качестве лидеров следует назвать Республику Казахстан и Кыргызскую Республику, в которых

разработаны и опубликованы проекты цифровых кодексов, предусматривающих регулирование ИИ.

2. Регулирование ИИ развивается в основном в направлении стимулирующих мер; в научных исследованиях значительное внимание уделяется гражданско-правовым вопросам, в частности интеллектуальной собственности.

3. В рамках публично-правовых вопросов наиболее обсуждаемой является проблема защиты прав граждан при использовании ИИ, включая защиту персональных данных в системах ИИ.

4. Отсутствует единое понимание базового термина «искусственный интеллект», различные страны демонстрируют значительное разнообразие подходов в данном вопросе. В законодательстве большинства государств отсутствует официальное определение ИИ. При этом определение, закрепленное в законодательстве Российской Федерации, представляется малоподходящим для практического применения (критерий «имитация когнитивной деятельности человека» слишком сложный и нечеткий для большинства участников регулируемых отношений). Таким образом, в части понятийного аппарата представляется более оправданным опираться на концепцию ОЭСР – ЮНЕСКО – ЕС.

Предложения

В основу регулирования ИИ в целях обеспечения национальной безопасности в государствах – членах ОДКБ рекомендуется заложить следующие принципы.

Принцип законности. Согласно данному принципу все участники отношений в области создания и использования технологий ИИ должны соблюдать нормы национального и международного права.

Принцип уважения и защиты прав и свобод человека. При создании и использовании технологий ИИ должны обеспечиваться права и свободы человека и гражданина, закрепленные в конституциях государств-членов, а также в международных соглашениях и договорах.

Принцип баланса развития и безопасности. Государства поощряют, стимулируют и создают условия для развития технологий ИИ таким образом, чтобы избежать существенных угроз для безопасности личности, общества и государства, а также для международной безопасности. Вместе с тем избыточное регулирование способно серьезно затормозить развитие технологий ИИ, в связи с чем применяется риск-ориентированный подход, ранжирование видов и степеней риска для жизненно важных интересов личности, общества и государства, а также ОДКБ в целом; для высокорисковых технологий приоритетным является принцип безопасности, для технологий с низким уровнем риска – принцип стимулирования развития.

Принцип государственного контроля. За созданием, внедрением и эксплуатацией систем ИИ, создающих или могущих создать высокие риски

причинения вреда интересам личности, общества и государства, требуется осуществлять государственный контроль.

Принцип обмена информацией. Следует создать механизм и каналы оперативного обмена информацией (в том числе конфиденциальной), необходимой для деятельности государств-членов в области создания ИИ и робототехники в целях обеспечения национальной безопасности (алгоритм уведомлений, участвующие структуры, формы документов, состав данных и пр.).

Принцип прозрачности и понятности. В системах ИИ принципы и результаты работы системы, во-первых, должны быть задокументированы, во-вторых, соответствующие сведения должны быть доступны заинтересованным лицам. Государства, организации и граждане должны быть осведомлены о том, почему получаются те или иные выходные результаты. Принцип прозрачности распространяется также на источники данных. Одновременно необходимо обеспечить баланс между прозрачностью и защитой конфиденциальной информации (коммерческая тайна, персональные данные и др.). В части генеративного ИИ результаты, полученные с использованием этих технологий, в обязательном порядке должны иметь соответствующую маркировку.

Принцип человеческого контроля. Должна быть закреплена возможность человека вмешаться в деятельность системы ИИ и остановить ее работу.

Принцип непрерывности. С учетом возможности выявления в процессе внедрения и эксплуатации у систем ИИ незапланированных, недокументированных свойств необходимо создать систему постоянного мониторинга инцидентов в сфере информационной безопасности, связанных с разработкой и эксплуатацией систем ИИ.

Принцип технологического суверенитета. Государствам – членам ОДКБ рекомендуется обеспечить необходимый уровень самостоятельности в области ИИ, в том числе посредством преимущественного использования технологий и решений, разработанных этими странами.

Принцип доступности. Технологии ИИ могут использоваться свободно в различных отраслях экономики и социальной сферы (за исключением государственного и муниципального управления и оборонно-промышленного комплекса).

Представляется целесообразной разработка ***модельной стратегии развития ИИ*** для обеспечения национальной безопасности, которая позволит гармонизировать подходы в политике государств – членов ОДКБ в сфере ИИ.

Данная модельная стратегия могла бы предусматривать:

1) гармонизацию подходов к пониманию ИИ, выработку общего согласованного понятийного аппарата;

2) дифференциацию подходов к правовому регулированию отношений, касающихся ИИ гражданского назначения, и отношений, касающихся ИИ

военного и двойного назначения, а также систем ИИ, применяемых непосредственно в целях обеспечения национальной безопасности;

3) введение для систем ИИ военного, двойного и специального назначения обязательного лицензирования деятельности по их созданию, внедрению и эксплуатации;

4) классификацию систем ИИ гражданского назначения в зависимости от характера и степени угроз, возникающих при их использовании (риск-ориентированный подход);

5) дифференцированный подход к правовому регулированию ИИ гражданского назначения в зависимости от степени риска, оценку эффективности правового регулирования с целью недопущения избыточных правовых барьеров для создания технологий ИИ;

6) формирование системы контроля за созданием и эксплуатацией систем ИИ с высоким риском, включая информирование уполномоченного государственного органа о разработке и эксплуатации высокорисковых систем ИИ;

7) разработку требований по безопасности систем ИИ, стандартов безопасности ИИ, механизмов сертификации систем ИИ на предмет соответствия указанным требованиям и стандартам;

8) создание единого кодекса этики в сфере ИИ на пространстве ОДКБ, в том числе путем присоединения к российскому Кодексу этики в сфере искусственного интеллекта (открыт для присоединения с 26 октября 2021 года);

9) создание в государствах-членах уполномоченного органа (системы органов) по ИИ с наделением его (их) следующими полномочиями:

– контроль за деятельностью лиц, ведущих работы в сфере ИИ;

– ведение учета проектов в сфере ИИ и обеспечение доступности информации о таких проектах (ведение информационных систем, библиотек, депозитариев);

– лицензирование деятельности по созданию систем ИИ военного и специального назначения;

– организация системы оценки соответствия систем ИИ установленным требованиям, в первую очередь организация разработки с учетом мирового опыта системы стандартов безопасности ИИ и методов оценки соответствия систем ИИ таким стандартам;

– в области международного сотрудничества – обмен информацией о работах в сфере ИИ, о возникающих угрозах, связанных с созданием и эксплуатацией систем ИИ.

Помимо вышеуказанного, также рекомендуется разработка организационно-правового механизма взаимодействия государств – членов ОДКБ в сфере ИИ, используемого в целях обеспечения национальной безопасности.