

КОММЕНТАРИЙ К РАМОЧНОЙ КОНВЕНЦИИ СОВЕТА ЕВРОПЫ ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ, ПРАВАХ ЧЕЛОВЕКА, ДЕМОКРАТИИ И ВЕРХОВЕНСТВЕ ПРАВА

МАРТЫНОВА Е. А., СУШКОВ С. П., ЕВСЕЕВ А. П., ШТОДИНА Д. Д.

Мартынова Екатерина Александровна — преподаватель департамента международного права факультета права Национального исследовательского университета «Высшая школа экономики», Москва, Россия (eamartynova@hse.ru). ORCID: 0000-0002-8995-4462.

Сушков Сергей Павлович — аспирант Аспирантской школы по праву Национального исследовательского университета «Высшая школа экономики», Москва, Россия (sushkov.mels@gmail.com). ORCID: 0000-0002-5366-4113.

Евсеев Артем Павлович — аспирант Аспирантской школы по праву Национального исследовательского университета «Высшая школа экономики», руководитель практики интеллектуальной собственности, IT и защиты информации Savina Legal, Москва, Россия (apevseev@hse.ru). ORCID: 0009-0006-8222-8077.

Штодина Дарья Дмитриевна — кандидат юридических наук, научный сотрудник департамента международного права факультета права Национального исследовательского университета «Высшая школа экономики», Москва, Россия (dashashtodina96@gmail.com). ORCID: 0000-0003-4730-9614.

Аннотация

Эта статья — одна из первых попыток дать научный комментарий к Рамочной конвенции Совета Европы об искусственном интеллекте, правах человека, демократии и верховенстве права 2024 года (далее — Конвенция). Авторы анализируют как общие характеристики Конвенции — ее юридическую природу, объект и цель — так и специальные вопросы, касающиеся реализации закрепленных в этом международном договоре принципов демократии и верховенства права, порядка использования средств правовой защиты, круга участников, а также механизмов имплементации Конвенции на национальном уровне. В статье также проводится сравнительно-правовой анализ этого договора и Акта Европейского союза об искусственном интеллекте, который до принятия Конвенции являлся единственным в своем роде документом, подробно регулирующим использование этой технологии. Авторы делают вывод о том, что Конвенция, хотя и воплощает «мягкую» модель международно-правового регулирования сферы искусственного интеллекта, тем не менее способна стимулировать развитие внутреннего законодательства государств-участников. Кроме того, некоторые из использованных в этом договоре подходов могут быть рассмотрены для целей разработки посвященных искусственному интеллекту международных правовых инструментов в рамках БРИКС, ШОС и/или ЕАЭС.

Ключевые слова

искусственный интеллект, рамочная конвенция, Совет Европы, права человека, верховенство права, риск-ориентированный подход

Для цитирования: Мартынова Е. А., Сушков С. П., Евсеев А. П., Штодина Д. Д. *Комментарий к Рамочной конвенции Совета Европы об искусственном интеллекте, правах человека, демократии и верховенстве права* // Журнал ВШЭ по международному праву (HSE University Journal of International Law). 2024. Т. 2. № 3. С. 52–63.

<https://doi.org/10.17323/jil.2024.23990>

Введение

Генеральный секретарь Совета Европы М. Пейчинович-Бурич сформулировала роль Рамочной конвенции об искусственном интеллекте, правах человека, демократии и верховенстве права (далее — Конвенция) следующим образом: «[Э]то первый в своем роде глобальный договор, который обеспечит соблюдение прав человека в области искусственного интеллекта. Это ответ на потребность в международном правовом стандарте, поддерживаемом государствами на разных континентах, которые разделяют одни и те же ценности, чтобы использовать преимущества искусственного интеллекта, одновременно снижая риски. С помощью этого нового договора мы стремимся обеспечить ответственное использование искусственного интеллекта, которое уважает права человека, верховенство права и демократию»¹. Х. Кольфсхутен и К. Скачар выделяют три основные

¹ Совет Европы принимает первый международный договор, касающийся искусственного интеллекта. URL: <https://www.coe.int/ru/web/portal/-/council-of-europe-adopts-first-international-treaty-on-artificial-intelligence#:~:text=Генеральн>

задачи, на решение которых направлена эта Конвенция: во-первых, устранение проблем в толковании прав человека в контексте применения искусственного интеллекта (далее — ИИ); во-вторых, необходимость правового закрепления основных прав человека применительно к ИИ; и, в-третьих, установление международных норм прав человека в области применения ИИ для развития международной торговли². Однако есть ли основания однозначно утверждать, что положения Конвенции представляют собой эффективный правовой инструмент решения этих задач? Действительно ли этот договор стоит в авангарде международно-правовых подходов к регулированию ИИ? Наконец, каково значение принятия Конвенции для Российской Федерации?

В статье сформулированы ответы на эти вопросы: в первой части дается комментарий к тексту Конвенции, во второй приводится сравнительный анализ этого международного договора и Акта об искусственном интеллекте Европейского союза 2024 года, и в третьей обсуждаются возможные последствия и значение принятия Конвенции для Российской Федерации.

1. Основные положения Конвенции и их имплементация на национальном уровне

1.1. Юридическая природа Конвенции, ее объект и цель

Конвенция представляет собой международный договор, целью которого выступает обеспечение соблюдения прав человека, верховенства права и правовых стандартов демократии на всех этапах проектирования, разработки и применения систем ИИ³. Развитие этих технологий приветствуется и поощряется государствами в связи с широкими возможностями, которые ИИ предоставляет для совершенствования иных технологий, роста промышленности и темпов торговли. Однако как на национальном, так и на международном уровне политическое, социальное и экономическое воздействие ИИ на различные общественные отношения охватывает проблемы, выходящие за рамки правового режима регулирования ИИ исключительно как технологии⁴. Принятие данной Конвенции как раз призвано ввести рамочное нормативно-правовое регулирование, отвечающее новым вызовам, встающим перед международным сообществом и отдельными государствами в связи с развитием ИИ, в частности, в сфере функционирования институтов демократии⁵, защиты прав и свобод⁶, преодоления социального неравенства и дискриминации, порождаемых использованием соответствующих компьютерных программ⁷.

Конвенция раскрывает термин «система ИИ»: это «машинная система, которая по явным или неявным целям определяет на основе полученных входных данных, как генерировать выходные данные, включая прогнозы, контент, рекомендации или решения, которые могут влиять на физическую или виртуальную среду», при этом корректно отмечается, что «системы искусственного интеллекта различаются по уровню автономности и адаптивности после развертывания»⁸ (здесь и далее перевод наш. — Авторы). Иными словами, в Конвенции воспроизведен так называемый «широкий» подход к определению системы ИИ, основанный на способности систем к самообучению и генерированию контента (в отличие от «узкого» подхода, основанного на способности системы ИИ решать конкретную прикладную задачу)⁹.

ый%20секретарь%20Совета%20Европы%20Мария, основан%20на%20уважении%20прав%20человека (дата обращения: 15.08.2024).

² Kolschooten H. v., Shachar C. *The Council of Europe's AI Convention (2023–2024): Promises and Pitfalls for Health Protection // Health Policy*. 2023. Vol. 138. P. 1–5.

³ См. ст. 1(1) Конвенции.

⁴ Crawford K. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Heaven : Yale University Press, 2021. P. 185–186.

⁵ Nemitz P. *Constitutional Democracy and Technology in the Age of Artificial Intelligence // Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2018. Vol. 376. № 2133. URL: <https://doi.org/10.1098/rsta.2018.0089> (дата обращения: 08.09.2024).

⁶ Donahoe E., Metzger M. M. *Artificial Intelligence and Human Rights // Journal of Democracy*. 2019. Vol. 30. № 2. P. 115–126.

⁷ Eubanks V. *Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor*. New York : St. Martin's Press, 2018. URL: https://openlibrary.org/books/OL26681102M/Automating_Inequality (дата обращения: 25.10.2024).

⁸ См. ст. 2 Конвенции.

⁹ Несмотря на повсеместный характер дискуссий об ИИ в последнее время, не существует единого «официального» определения ИИ. В некоторых случаях технические описания, предлагаемые специалистами в области компьютерных наук, не подходят для юридического анализа, например, когда ИИ определяется как алгоритм, что, в свою очередь, требует отдельного определения юридического содержания соответствующих терминов. Подробнее о широком и узком подходах к определению систем ИИ см., например: Lee J. *Artificial Intelligence and International Law*. Springer, 2022. P. 6–7.

Приведенное в Конвенции определение не содержит указания на возможное двойное назначение (военное и гражданское) систем ИИ. Вместе с тем важно отметить, что из сферы действия Конвенции исключена деятельность, связанная с национальной обороной¹⁰. Таким образом, обязательства участников Конвенции по обеспечению прозрачности, подотчетности и ответственности за возможные неблагоприятные последствия не распространяются на действия, связанные с разработкой и применением этой технологии в военных целях. Однако применение ИИ в оборонной сфере в качестве автономного летального оружия способно серьезно повлиять на геополитический баланс между государствами, создавая новые международные асимметрии¹¹.

С точки зрения объема и сфер применения Конвенции предполагается, что государства-участники примут законодательные, административные и/или иные меры, необходимые для соблюдения ее положений как государственными органами, так и частными субъектами, действующими от их имени¹². Для регулирования деятельности частных лиц, действующих не от имени государства, Конвенцией предусмотрена альтернатива: стороны могут распространить принципы и обязательства, изложенные в Конвенции, на частный сектор (уравнять его в регулировании с государственным) или принять другие меры для управления рисками применения систем ИИ частными лицами таким образом, чтобы это соответствовало объекту и цели данного международного договора¹³. О выбранном способе исполнения обязательства по регулированию деятельности частного сектора необходимо сообщить при подписании или при сдаче на хранение документа о ратификации, при принятии, утверждении Конвенции или присоединении к ней (избранный способ можно в дальнейшем изменить). При этом не допускается отступление или ограничение применения стороной своих международных обязательств в отношении прав человека, демократии и верховенства права¹⁴. Такая «развилка» в методах регулирования деятельности частного сектора представляется не вполне удачной, так как создает существенную асимметрию в объеме обязательств государств — участников Конвенции в зависимости от выбранного варианта.

1.2. Демократия и верховенство права

В то время как Конвенция посвящена взаимодействию «искусственного интеллекта, прав человека, демократии и верховенства права», только статья 5 этого договора прямо регулирует потенциальное воздействие систем ИИ на демократические процессы и принцип верховенства права. Во-первых, согласно статье 5, государства обязаны принимать и поддерживать меры, направленные на недопущение использования ИИ с целью нарушения целостности, независимости и эффективности демократических институтов и процессов. Конвенцией предусмотрена защита ряда основополагающих демократических принципов, в частности разделения властей, независимости судебной власти и доступа к правосудию. Во-вторых, государства обязаны принимать и поддерживать меры, направленные на защиту демократических процессов при разработке систем ИИ, в том числе таких аспектов демократического процесса, как возможность граждан на равных началах участвовать в публичной дискуссии, а также свободно формировать собственное мнение.

В Конвенции не проведено четкого различия между правилами, ориентированными на защиту демократических процессов и институтов, и правилами, ориентированными на защиту верховенства права. Напротив, в статье 5 договора эти две категории рассматриваются в их взаимной связи. В качестве одного из проявлений защиты демократических институтов указывается обеспечение независимости судебной власти, а также еще одно проявление принципа верховенства права: доступ к правосудию. Поскольку верховенство права представляет один из признаков демократического режима, любая мера по защите верховенства права также будет способствовать защите демократических институтов. Следовательно, исходя из текста статьи 5 Конвенции, предмет регулирования Конвенции стоит рассматривать, с одной стороны, как регулирование систем ИИ в свете прав человека, а с другой — как регулирование соотношения ИИ с принципом верховенства права и демократическими процессами и институтами.

¹⁰ См. ст. 3(4) Конвенции.

¹¹ Johnson J. *Artificial Intelligence and Future Warfare: Implications for International Security* // *Defense & Security Analysis*. 2019. Vol. 35. № 2. P. 147–169.

¹² См. ст. 1(2) и 3(1)(a) Конвенции.

¹³ См. ст. 3(1)(b) Конвенции.

¹⁴ См. там же.

В пояснительной записке к Конвенции, которая не является официальным источником ее толкования, но позволяет лучше уяснить идеи, стоящие за положениями этого договора¹⁵, перечислены возможные способы негативного воздействия систем ИИ на демократию и верховенство права. Среди них — возможность использования для распространения информации, не соответствующей действительности, в том числе контента, сформированного системой ИИ (так называемые дипфейки) или для установления незаконной или произвольной слежки за гражданами¹⁶. Данные способы использования ИИ могут нанести ущерб избирательному процессу, ограничить свободу слова и собраний, подорвать доверие между гражданами и государством, привести к дискриминационным практикам, а также повлиять на независимость и беспристрастность судей и эффективность судебной системы в целом.

Стоит отметить, что в Конвенции не конкретизирована территориальная сфера действия обязательств, предусмотренных статьей 5. Правом международных договоров не предусмотрена ни презумпция того, что международный договор может устанавливать экстерриториальные обязательства государства, ни презумпция обратного¹⁷. Возможность экстерриториального применения обязательств, вытекающих из международного договора, можно установить только исходя из текста, объекта и цели каждого конкретного договора¹⁸. Например, Международный пакт об экономических, социальных и культурных правах 1966 года¹⁹, Конвенция о ликвидации всех форм дискриминации в отношении женщин 1979 года²⁰ и Конвенция о правах инвалидов 2006 года²¹, как и данная Конвенция, не устанавливают территориального предела действия обязательств государств-участников. Это позволяет утверждать, что государства должны соблюдать обязательства по данным международным договорам, даже действуя за пределами собственной территории²². Международный Суд отмечал, что положения Конвенции о ликвидации всех форм расовой дискриминации, также не содержащие ограничений *ratione loci*, применяются экстерриториально, «как и другие положения документов такого рода»²³ (курсив наш. — Авторы), то есть международных договоров в области защиты прав человека. Поскольку Конвенция декларирует широкую цель обеспечения соблюдения прав человека, верховенства права и правовых стандартов демократии на всех этапах проектирования, разработки и применения систем ИИ²⁴, веские доводы против распространения действия Конвенции за пределы территории государств-участников, на первый взгляд, отсутствуют.

Следовательно, допустимо толкование, согласно которому государства-участники должны регулировать не только возможное влияние систем ИИ на демократические институты, процессы и верховенство права в пределах своей юрисдикции, но и на потенциальное воздействие ИИ на демократические институты, процессы и верховенство права в других государствах. Соответственно, поскольку использование данной технологии может оказать существенное влияние на избирательный процесс, ограничение воздействия ИИ на демократические институты и процессы должно способствовать укреплению принципа невмешательства во внутренние дела государства²⁵.

¹⁵ *Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, CETS 225, 5.IX.2024 (далее — Explanatory Report), Preamble, Part II.

¹⁶ Explanatory Report, § 43.

¹⁷ Milanovic M. *Extraterritorial Application of Human Rights Treaties: Law, Principles, and Policy*. Oxford University Press, 2011. P. 10.

¹⁸ Ibid. P. 11.

¹⁹ Международный пакт об экономических, социальных и культурных правах от 16 декабря 1966 года // Ведомости Верховного Совета СССР. 1976. № 17(1831).

²⁰ Конвенция о ликвидации всех форм дискриминации в отношении женщин от 18 декабря 1979 года // Ведомости Верховного Совета СССР. 1982. № 25 Ст. 464.

²¹ Конвенция о правах инвалидов от 13 декабря 2006 года // Собрание законодательства РФ. 11 февраля 2013. № 6. Ст. 468.

²² Janig P. *Extraterritorial Application of Human Rights* // *Elgar Encyclopedia of Human Rights*. Vol 2 / ed. by C. Binder, M. Nowak, J. Hofbauer. P. Jang. Edward Elgar Publishing, 2022. P. 182–183; Milanovic M. *Op.cit.* P. 17.

²³ Application of the International Convention on the Elimination of all Forms of Racial Discrimination (*Georgia v. Russian Federation*), Provisional Measures. Order of 15 October 2008. I.C.J. Reports, 2008. P. 353, § 109.

²⁴ См. Ст. 1(1) Конвенции.

²⁵ Декларация о принципах международного права, касающихся дружественных отношений и сотрудничества между государствами в соответствии с Уставом Организации Объединенных Наций, принята резолюцией 2625 (XXV) Генеральной Ассамблеи ООН от 24 октября 1970 года.

1.3. Основные обязательства государств — участников Конвенции

Несмотря на перечисление в Конвенции обязательств государств-участников, использование ряда «защитных оговорок» (англ.: *saving clauses*) предопределяет ее рамочный характер. Так, в качестве общего принципа регулирования предусмотрено, что участники исполняют свои обязательства по Конвенции таким образом, который будет соответствовать их национальной правовой системе (англ.: *in a manner appropriate to its domestic legal system*)²⁶. По тексту Конвенции обязательства государств сформулированы скорее как «мягкие» цели и обязательства поведения (англ.: *obligation of conduct*), нежели обязательства результата (англ.: *obligation of result*): «должны принимать или сохранять меры, направленные на...», «должны принять меры, когда это будет уместно, для...», «должны стремиться обеспечить, чтобы в той мере, в которой это уместно в конкретных обстоятельствах», «призываются к созданию условий в той мере, в какой это уместно, для...», «в той мере, в какой средства правовой защиты требуются в соответствии с международными обязательствами и в соответствии с внутренней правовой системой, должны принимать или сохранять меры для обеспечения», «должны оценить необходимость» и т. п.

Такой подход, с одной стороны, обеспечивает гибкость применения Конвенции и, вероятно, будет способствовать расширению круга ее участников, а с другой — «размывает» содержание обязательств участников и резервирует обширное пространство для их толкования. При этом возможное смещение баланса между нормативностью и определенностью, о котором писал М. Коскенниemi²⁷, в сторону нормативности неизбежно приведет к непоследовательности практики и, в свою очередь, к «политизации» соответствующего регулирования. В частности, предусмотренное Конвенцией обязательство сторон принять меры, направленные на защиту демократических процессов в рамках жизненного цикла систем ИИ, включая обеспечение возможности для граждан свободно формировать свое мнение²⁸, может быть диаметрально противоположно реализовано в государствах, в том числе в зависимости от выбранного подхода к регулированию деятельности социальных сетей.

Конвенция устанавливает обязательство участников внедрить адекватные требования прозрачности и надзора (англ.: *adequate transparency and oversight requirements*) в отношении деятельности в рамках жизненного цикла систем ИИ, охватывающего и вопросы идентификации создаваемого такими системами контента, с учетом конкретных условий и рисков²⁹. С этим обязательством связано требование обеспечить подотчетность и ответственность за возможные неблагоприятные последствия использования систем ИИ для прав человека, демократии и верховенства права³⁰.

Государства — участники Конвенции должны также принять меры к тому, чтобы в системах ИИ уважалось равенство, включая гендерное равенство, соблюдался запрет дискриминации и не нарушалось право на неприкосновенность частной жизни³¹. При этом в соответствующих статьях данного международного договора предусмотрены положения о том, что такие обязательства реализуются государствами с учетом норм международного и национального права. Представляется, что реальное «наполнение» данных обязательств в государствах различных правовых систем может значительно отличаться, в частности, в отношении обеспечения гендерного равенства и подходов к основаниям допустимого ограничения права на уважение частной жизни. В целом, можно предположить, что выбор и «калибровка» инструментов, заложенных в Конвенции, и их имплементация в национальном регулировании ИИ будут предопределяться особенностями политического режима государства — участника Конвенции: в частности, степенью вовлеченности заинтересованных сторон (стейкхолдеров) в процесс формирования нормативного регулирования, эффективностью институтов, определяющих правила поведения участников жизненного цикла систем ИИ, а также ролью гражданского общества и организаций по защите прав и свобод человека³².

²⁶ См. ст. 6 Конвенции.

²⁷ Koskenniemi M. *From Apology to Utopia: The Structure of International Legal Argument*. Cambridge : Cambridge University Press, 2006.

²⁸ См. ст. 5(2) Конвенции.

²⁹ См. ст. 8 Конвенции.

³⁰ См. ст. 9 Конвенции.

³¹ См. ст. 10 и 11 Конвенции.

³² Обзор режимов национальной политики в области ИИ и их типологии в зависимости от политического режима см.: Filgueiras F. *Artificial Intelligence Policy Regimes: Comparing Politics and Policy to National Strategies for Artificial Intelligence // Global Perspectives*. 2022. Vol. 3. № 1. P. 323–362. URL: <https://doi.org/10.1525/gp.2022.32362> (дата обращения: 08.09.2024).

1.4. Средства правовой защиты

Участники Конвенции обязаны обеспечить доступность средств правовой защиты лицам, чьи права человека были нарушены в связи с использованием систем ИИ (опять же, с оговоркой, что такие меры принимаются в той степени, в которой они соответствуют требованиям национальной правовой системы)³³. В качестве базовых процессуальных гарантий защиты прав человека при взаимодействии с системами ИИ Конвенцией предусмотрены: уведомление любых лиц о вступлении их во взаимодействие с системами ИИ, документирование информации об использовании систем ИИ, потенциально нарушающих права человека, и возможность для заинтересованных лиц получить доступ к такой информации и подать жалобу в компетентный государственный орган³⁴. Однако какой именно подход будет применяться, чтобы определить, были ли нарушены права заявителей, остается за рамками этого международного договора, что оставляет пространство для применения самых разных моделей. Таким образом, Конвенция лишь задает общий процессуальный вектор, гарантируя наличие некоего средства правовой защиты.

1.5. Исключения в целях обеспечения национальной безопасности и научных исследований

Стороны Конвенции освобождаются от соблюдения принятых на себя обязательств при осуществлении деятельности, связанной с защитой интересов национальной безопасности, при условии соблюдения норм международного права, в том числе международных обязательств в области прав человека, и при уважении демократических институтов и процессов³⁵. Мотивы введения такого исключения понятны, однако опасность состоит в том, что государства могут применять его расширительно, не предоставляя пояснений о причинах применения данного исключения и оправдывая это тем, что само пояснение причин создает угрозу для национальной безопасности. Конвенция также не ограничивает участников в проведении научно-исследовательских работ в области ИИ при условии, что тестирование систем ИИ или иные работы не связаны с рисками в отношении прав человека, демократии и верховенства права³⁶.

1.6. Круг участников Конвенции

Конвенция открыта для подписания государствами — членами Совета Европы, Европейского союза, а также государствами, участвовавшими в ее разработке. Вероятно, имеются в виду государства, представители которых принимали участие в работе Комитета по искусственному интеллекту: Аргентина, Австралия, Канада, Коста-Рика, Ватикан, Израиль, Япония, Мексика, Перу, США и Уругвай. Подписание прошло в Вильнюсе (Литва) 5 сентября 2024 года в ходе конференции министров юстиции. Конвенцию подписали Андорра, Грузия, Исландия, Норвегия, Республика Молдова, Сан-Марино, Соединенное Королевство, Израиль, США, а также Европейский союз³⁷. После вступления Конвенции в силу (в первый день месяца, следующего за истечением трехмесячного срока после даты ратификации Конвенции пятью подписавшими сторонами, включая не менее трех государств — членом Совета Европы) присоединиться к ней смогут государства, не участвовавшие в ее разработке, при условии, что такое присоединение будет одобрено решением, принятым большинством в соответствии со статьей 20(d) Устава Совета Европы, и единогласным голосованием представителей сторон Конвенции, имеющих право заседать в Комитете Министров этой международной организации. Строгий порядок присоединения к Конвенции не уникален для договоров Совета Европы: большинство из них, включая так называемые «открытые» договоры, то есть допускающие присоединение государств, не входящих в Совет Европы, требуют единогласного согласия сторон³⁸.

³³ См. ст. 14(1) Конвенции.

³⁴ См. ст. 14(2) Конвенции.

³⁵ См. ст. 3(2) Конвенции.

³⁶ См. ст. 3(3) Конвенции.

³⁷ Council of Europe. *Council of Europe opens first ever global treaty on AI for signature*. URL: <https://www.coe.int/en/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature> (дата обращения: 08.09.2024).

³⁸ Council of Europe. *Participation of Non-member States*. URL: <https://www.coe.int/ru/web/conventions/participation-of-non-member-states> (дата обращения: 13.10.2024).

1.7. Имплементация положений Конвенции

Как было отмечено выше, рамочный характер Конвенции оказал влияние на формулировки обязательств, принимаемых на себя государствами. В тексте не установлено жестких требований по принятию конкретных мер. Отсюда положения этого международного договора являются по большей части несоизмеримыми, а характер обязательств, установленных Конвенцией, оставляет за государствами широкие дискреционные полномочия по ее имплементации. Помимо этого, широкая свобода усмотрения в имплементации выражается в праве государств-участников самим определять, в каком объеме этот договор будет распространяться на разработку и использование систем ИИ в частном секторе³⁹.

Конвенцией предусмотрены два механизма содействия государствам в исполнении взятых на себя обязательств. Во-первых, статьей 23 учреждается Конференция государств-участников, которая не является постоянным органом и созывается по мере необходимости Генеральным Секретарем Совета Европы, или если большинство государств-членов или Комитет Министров Совета Европы потребуют созыва Конференции. Как отмечено в пояснительной записке к Конвенции, отсутствие правила о постоянной работе Конференции отражает «гибкость механизма последующего контроля, установленного данной Рамочной Конвенцией»⁴⁰.

Во-вторых, статьей 24 установлена обязанность предоставлять Конференции государств-участников отчет о действиях по имплементации Конвенции. Данный отчет должен быть предоставлен в течение первых двух лет после того, как государство стало стороной Конвенции, а затем отчеты предоставляются периодически. Формат и процедура подготовки отчетов будут установлены Конференцией государств-участников.

Кроме того, статьей 26 Конвенции предусмотрено обязательство создать эффективные внутренние механизмы для надзора за соблюдением обязательств, предусмотренных настоящей Конвенцией. В пояснительной записке к Конвенции указано, что данное обязательство может быть реализовано или посредством учреждения отдельного государственного органа, или посредством передачи уже существующему государственному органу полномочий по контролю за исполнением Конвенции⁴¹. При этом, как следует из п. 2 статьи 26, данные органы должны действовать «независимо и беспристрастно», им должны быть предоставлены необходимые полномочия и ресурсы для эффективности исполнения их функций по контролю за соблюдением Конвенции.

2. Конвенция и Акт Европейского союза об искусственном интеллекте: сравнительный анализ

В отличие от Конвенции, Акт Европейского союза (далее — ЕС) об искусственном интеллекте⁴² (далее — Акт) предусматривает более детальное регулирование ИИ. Принятие Акта к тому же явилось логичным продолжением правовой политики ЕС: в 2016 году был принят Общий регламент защиты данных (англ.: *General Data Protection Regulation*)⁴³, вступивший в силу в мае 2018 года. Данный регламент установил единые правила сбора и последующего оборота персональных данных в масштабах всего ЕС. В свою очередь, Акт преследует достижение такой же цели, но применительно к информации, обрабатываемой ИИ-системами. В начале сравнительного анализа Акта и Конвенции необходимо отметить, что Акт является актом прямого действия, создающим конкретные права и обязанности для участников рынка на территории всего ЕС, в то время как Конвенция направлена на стимулирование развития национального регулирования государств-участников, а потому выстраивает лишь общие регуляторные контуры и принципы.

Акт принят в форме регламента и, соответственно, действует непосредственно на территории всех государств — членов ЕС. Цели Акта шире, чем цели Конвенции, и включают не только защиту

³⁹ См. Ст. 3(1)(b) Конвенции. См. более подробно об этом в разделе 1.1 выше.

⁴⁰ Explanatory Report, § 131.

⁴¹ Explanatory Report, § 141–142.

⁴² 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. PE/24/2024/REV/1 // OJ L, 2024/1689, 12.7.2024.

⁴³ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. PE 17 2016 INIT OJ : JOL_2016_119_R_0001.

демократии, прав человека и принципа верховенства права, но и сохранение единого рынка ЕС на фоне распространения товаров и услуг, использующих ИИ. Конечная цель Акта — создать человеко-ориентированный ИИ с учетом необходимости защиты здоровья, безопасности и основных прав человека, а также предотвращения негативных последствий от использования ИИ-технологий при сохранении инновационного развития.

Данные цели логичным образом вытекают из пояснительного меморандума Еврокомиссии, подготовленного к первоначальному проекту Акта: ЕС должен придерживаться сбалансированного подхода при выстраивании нормативного регулирования ИИ-систем⁴⁴. С одной стороны, использование ИИ-технологий улучшает прогнозирование, оптимизирует операции и распределение ресурсов, а также персонализирует предоставление услуг, иным образом позволяет достигать положительных результатов для общества, компаний и экономики. Однако, с другой стороны, те же элементы и методы, которые являются преимуществами ИИ, могут обернуться новыми рисками или негативными последствиями для отдельных лиц или общества в целом. Поэтому задача нормативного регулирования состоит в том, чтобы граждане ЕС могли получать выгоду от новых технологий, при этом такие технологии должны быть разработаны и функционировать на основании ценностей ЕС, основных прав человека и гражданина.

Восприятие ИИ через призму прав человека выступает очередным свидетельством продолжающейся «конституционализации» права ЕС, которое превратилось в систему норм прямого действия, дающих осязаемые гарантии для всех граждан этой наднациональной организации⁴⁵. Именно поэтому, как отмечают ученые, несмотря на преследование целей поддержания верховенства права и демократии, Акт стоит на защите, прежде всего, прав человека, общественной безопасности и здравоохранения⁴⁶.

По сравнению с Конвенцией концептуальная проработка Акта ЕС началась гораздо раньше. Так, в 2019 году под эгидой Еврокомиссии были сформулированы семь постулатов, на основании которых должно строиться правовое регулирование ИИ-систем, а именно: (1) управление и контроль со стороны человека; (2) обеспечение технологического развития и сохранение безопасности; (3) защита частной жизни и персональных данных; (4) открытость; (5) плюрализм, запрет дискриминации и равенство; (6) благополучие общества и окружающей среды; а также (7) ответственность⁴⁷.

Кроме того, в отличие от Конвенции, Акт прямо определяет сферу своего действия, будучи распространен в соответствии со статьей 2(1) экстерриториально на лиц, которые разработали или владеют ИИ-системами и выводят их на рынок ЕС или вводят в эксплуатацию под своим именем, независимо от взимания платы; на пользователей ИИ-систем (кроме тех, кто использует их для личных нужд), находящихся в ЕС; на дистрибьюторов и импортеров ИИ-систем при их ввозе и дистрибуции в ЕС; на производителей ИИ-систем, вне зависимости от их местоположения; а также на любых лиц, находящихся в ЕС, на которых воздействует ИИ.

В отличие от Конвенции, которая содержит достаточно общие положения о необходимости соблюдения прав человека, статья 5 Акта ЕС прямо запрещает несколько «практик» использования ИИ. Среди них:

- (а) вывод на рынок, ввод в эксплуатацию или использование ИИ, который применяет подсознательные, манипулятивные или обманные методы;
- (б) использование ИИ, эксплуатирующего уязвимости людей, связанные с возрастом, инвалидностью или социально-экономической ситуацией;
- (с) применение ИИ для оценки или классификации людей на основе их социального поведения или личных характеристик;
- (d) использование ИИ для оценки риска совершения преступления на основе профиля или личных качеств человека, кроме случаев, когда ИИ поддерживает оценку, основанную на объективных фактах, установленных людьми;

⁴⁴ Explanatory Memorandum to Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending certain union legislative acts. COM(2021) 206 final 2021/0106 (COD), URL: <https://artificialintelligenceact.eu/wp-content/uploads/2022/05/AIA-COM-Proposal-21-April-21.pdf> (дата обращения: 21.10.2024).

⁴⁵ Rittberger B., Schimmelfennig F. *The Constitutionalization of the European Union*. Routledge, 2007. P. 7.

⁴⁶ Ho C. W.-L., Caals K. *How the EU AI Act Seeks to Establish an Epistemic Environment of Trust* // *Asian Bioethics Review*. 2024. Vol. 16. P. 345–372, 346.

⁴⁷ High-Level Expert Group on Artificial Intelligence. *Ethics Guidelines for Trustworthy AI*. 8 April 2019. P. 14–20; *Building Trust in Human-Centric Artificial Intelligence* (Commission Communication). COM(2019) 168 final. Brussels, 8 April 2019. P. 3–6.

(e) создание или расширение баз данных распознавания лиц через нецеленаправленное извлечение изображений из интернета или камер видеонаблюдения;

(f) использование ИИ для определения эмоций на рабочем месте или в учебной среде, за исключением медицины и безопасности;

(g) применение биометрических систем для категоризации людей по биометрическим данным с целью установления их расы, политических взглядов, религиозных убеждений, сексуальной ориентации и т.п., кроме случаев обеспечения безопасности;

(h) использование систем биометрической идентификации в реальном времени в общественных местах для целей правоприменения, кроме случаев предотвращения серьезных угроз, поиска пропавших лиц или преступников.

В отличие от Конвенции, Акт разделяет все виды ИИ-систем на три категории: высокорисковые ИИ-системы, общие языковые модели, а также все остальные системы ИИ. Акт приводит примерный перечень высокорисковых ИИ-систем, и этот перечень может обновляться Еврокомиссией. К таким системам относятся ИИ-системы, которые помогают в трудоустройстве, оценивают уровень знаний учащихся, анализируют эмоции человека, проводят биометрическую идентификацию и т.п. Владельцы таких ИИ-систем должны создать эффективный механизм управления рисками на всем жизненном цикле системы (от разработки до технической поддержки пользователей). Акт также требует, чтобы информация, на основании которой обучается ИИ-алгоритм, соответствовала стандартам качества (например, была репрезентативной). Кроме того, согласно Акту, работа высокорисковых ИИ-систем должна находиться под постоянным человеческим контролем, а информация об их работе (например, о точности результатов) должна быть доступна пользователям. В ряде случаев владельцам высокорисковых ИИ необходимо пройти техническую сертификацию. При выводе на рынок такой системы, ее владелец или его представитель должны подать заявление для включения в особый реестр. Для общих языковых моделей Акт устанавливает схожие правила, выделяя модели с системными рисками как отдельную категорию. Тем самым Акт вводит риск-ориентированное регулирование ИИ-систем в зависимости от их публичной значимости, той информации, которую они обрабатывают, степени тяжести последствий злоупотребления такими ИИ-системами и пр.⁴⁸

Акт также закрепляет институциональную модель регулирования и контроля за сферой ИИ в рамках ЕС. На основании решения Еврокомиссии создан отдельный орган — Управление по вопросам искусственного интеллекта (англ.: *AI Office*), которое будет заниматься стимулированием саморегулирования (например, созданием кодексов лучших практик)⁴⁹. Это управление является частью административной структуры Генерального директората по коммуникационным сетям, контенту и технологиям (статья 1 Акта). Статья 3(47) также закрепляет, что этот орган выполняет «функцию Еврокомиссии по содействию реализации, мониторингу и надзору за системами ИИ, моделями ИИ общего назначения и управлением ИИ». Таким образом, данное управление будет выступать, скорее всего, как одна из служб Еврокомиссии, а не в качестве независимого органа в рамках ЕС⁵⁰. В рамках своей компетенции *AI Office* обеспечивает исполнение и реализацию Акта, а также содействует в разработке политик в области ИИ в рамках ЕС и на международном уровне (статья 2 и статья 3(2) Акта).

Кроме того, будет создан Совет ЕС по вопросам ИИ, который займется выработкой общеевропейских подходов к регулированию ИИ, включая представителей всех государств — членов ЕС. Акт также устанавливает, что каждое государство — член ЕС должно создать свои национальные органы по сертификации высокорисковых ИИ-систем и контролю за их использованием.

Наряду с этим Акт, в отличие от Конвенции, закрепляет механизм юридической ответственности (см. статью 99 Акта). Так, согласно данной статье, государства-члены обязаны установить конкретные санкции за нарушения положений Акта, в том числе административные штрафы в пределах до 15 млн евро или до 3% размера общемировой выручки, а также предусмотреть основания для снижения сумм штрафов при наличии смягчающих обстоятельств. Как только компетентные национальные

⁴⁸ Novelli C., Casolari F., Rotolo A. et al. *AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act* // Digital Society. 2024. Vol. 3. Article № 13. P. 13–15.

⁴⁹ Commission Decision of 24.01.2024 establishing the European Artificial Intelligence Office. C(2024) 390 final.

⁵⁰ Lütz F. *The AI Act, Gender Equality and Non-Discrimination: What Role for the AI Office?* // ERA Forum. 2024. Vol. 25. P. 79–95. P. 87.

органы установят подобные санкции и механизм их наложения, государства-члены должны незамедлительно уведомить об этом Еврокомиссию.

На текущий момент именно Акт ЕС является одним из ярких примеров детального регулирования ИИ. Он направлен на сохранение общего рынка ЕС и защиту фундаментальных прав человека и гражданина. Основное внимание уделяется ИИ-системам, несущим наибольший риск для человека и общества. Введение процедуры сертификации ИИ-систем, регистрация их владельцев, создание институциональной модели регуляторов на уровне ЕС и национальном уровне, закрепление механизма юридической ответственности, а также создание экспертного представительства в виде Совета ЕС по вопросам ИИ формирует основу для гибкого развития нормативного регулирования ИИ в рамках ЕС.

Между тем использование основанной на правах человека модели как в Акте ЕС, так и в Конвенции, имеет один и тот же изъян: распространяя сферу своего применения в том числе и на частных акторов⁵¹, они не предусматривают конкретных материально-правовых норм, которые «трансформировали» бы содержание норм международного права прав человека, рассчитанных на государства как на адресатов соответствующих обязанностей. Существующие в этой области нормы базируются на соблюдении баланса между правами и свободами человека и легитимными целями, которые могут преследоваться государством, или же правами других людей, а также, как правило, содержат формулировки, допускающие значительное поле для свободы толкования самими государствами. Отсюда «язык», внутренняя структура и природа этих конструкций не позволяют без надлежащей трансформации и адаптации по аналогии применять их к частным акторам и, прежде всего — к коммерческим юридическим лицам⁵². Без необходимых изменений ссылка на применение или соблюдение прав человека будет оставаться не более чем декларацией⁵³.

3. Возможные последствия принятия Конвенции для Российской Федерации

Конвенция налагает обязательства только на государства, которые выразили согласие на ее обязательность, и не содержит каких-либо норм, распространяющихся на государства, не являющиеся ее участниками. Несмотря на то что Конвенция содержит декларативное положение о стремлении сторон способствовать тому, чтобы неприсоединившиеся государства действовали в соответствии с заложенными в ней принципами⁵⁴, это положение не создает каких-либо нормативных обязательств.

Вопрос о присоединении Российской Федерации к Конвенции едва ли будет стоять в ближайшее время, в том числе в связи с требующимся единогласным голосованием представителей сторон Конвенции, имеющих право заседать в Комитете министров, для присоединения государства, не являющегося членом Совета Европы. Вместе с тем двухлетний опыт работы Комитета по ИИ над текстом Конвенции и ее отдельные положения могут быть полезны при формировании российской позиции по вопросу нормативного регулирования деятельности с использованием систем ИИ. Важнейшими принципами деятельности государственных органов и частных лиц в рамках жизненного цикла систем ИИ представляются принципы прозрачности, надежности (англ.: *reliability*), анализа рисков, подотчетности и ответственности за негативные последствия. Гарантии в области информирования граждан, предусмотренные Конвенцией, также могут быть восприняты российской правовой системой. Например, в форме возможного маркирования контента, созданного системой ИИ, а также информирования о том, что осуществляется взаимодействие с системой ИИ (в частности, это возможно в случае, когда потребителю оказываются услуги посредством телефонного звонка). Наконец, через некоторое время анализ нового опыта государств — участников Конвенции в части обеспечения средств правовой защиты для граждан, чьи права нарушены при использовании систем ИИ, может быть полезен для совершенствования российского законодательства, особенно в областях, «чувствительных» с точки зрения защиты прав граждан, например, в сфере использования систем распознавания лиц.

⁵¹ Ст. 3 (1) (а) Конвенции; ст. 2 (1) Акта ЕС.

⁵² См.: Rusinova V., Korotkov S. *Mandatory Corporate Human Rights Due Diligence Models: Shooting Blanks?* // *Russian Law Journal*. 2021. Vol. 9. № 4. P. 43–45.

⁵³ Butler J. *The Corporate Keepers of International Law* // *American Journal of International Law*. 2020. Vol. 114. № 2. P. 189, 216–218.

⁵⁴ Ст. 25(1) Конвенции.

Заключение

В целом Конвенция воплощает «мягкую» модель международно-правового регулирования сферы искусственного интеллекта, оставляя за государствами существенные дискреционные полномочия и не предусматривая строгой системы контроля за ее исполнением. Это существенно отличает данный международный договор от принятого в ЕС Акта, который является примером первого детального урегулирования использования ИИ-систем на наднациональном уровне, а также имеет прямое действие, создавая конкретные права и обязанности для государства и участников рынка. При этом принятие правил, закрепленных в Конвенции, определенно будет являться положительным стимулом для развития внутреннего законодательства, регулирующего разработку и использование систем ИИ. В связи с этим представляется, что данная модель могла бы быть рассмотрена для целей разработки аналогичных международных правовых инструментов на площадках БРИКС, ШОС и/или ЕАЭС.

COMMENTARY ON THE COUNCIL OF EUROPE FRAMEWORK CONVENTION ON ARTIFICIAL INTELLIGENCE, HUMAN RIGHTS, DEMOCRACY AND THE RULE OF LAW

MARTYNOVA E., SUSHKOV S., EVSEEV A., SHTODINA D.

Ekaterina Martynova — Lecturer at the School of International Law, Faculty of Law, HSE University, Moscow, Russia (eamartynova@hse.ru).
ORCID: 0000-0002-8995-4462.

Sergey Sushkov — Doctoral Student, Doctoral School of Law, HSE University, Moscow, Russia (sushkov.mels@gmail.com).
ORCID: 0000-0002-5366-4113.

Artyom Evseev — Doctoral Student, Doctoral School of Law, HSE University; Head of Intellectual Property, IT and Information Protection Practice at Savina Legal, Moscow, Russia (apevseev@edu.hse.ru).
ORCID: 0009-0006-8222-8077.

Dariya Shtodina — Candidate of Sciences in Law, Research Associate, School of International Law, Faculty of Law, HSE University, Moscow, Russia (dashashtodina96@gmail.com).
ORCID: 0000-00034730-9614.

Abstract

The article is one of the first attempts to provide a scientific commentary on the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law of 2024 (hereinafter — the Convention). The authors analyse both general characteristics of the Convention — its legal nature, object and purpose — and specific issues concerning the implementation of the principles of democracy and the rule of law enshrined in this international treaty, the procedure for the use of legal remedies, the scope of parties, as well as the mechanisms of implementing the Convention at the national level. Furthermore, the article contains a comparative legal analysis of this treaty with the European Union Act on Artificial Intelligence, which, prior to the adoption of the Convention, constituted a unique exemplar of the most comprehensive international legal document regulating the utilisation of this technology. The authors conclude that, despite embodying a “soft” model of international legal regulation in the sphere of artificial intelligence, the Convention is nevertheless capable of stimulating the development of domestic legislation in the States Parties. Furthermore, some of the approaches used in this treaty could be considered for the development of international legal instruments dedicated to artificial intelligence within the BRICS, SCO and/or EAEU.

Keywords

artificial intelligence, framework convention, human rights, Council of Europe, rule of law, risk-based approach

Citation: Martynova E., Sushkov S., Evseev A., Shtodina D. *Kommentariy k Ramochnoy konventsii Soveta Evropy ob iskusstvennom intellekte, pravakh cheloveka, demokratii i verkhovenstve prava* [Commentary on the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law] // Zhurnal VSHÉ po mezhdunarodnomu pravu (HSE University Journal of International Law). 2024. Vol. 2. No. 3. P. 52–63. (In Russian).

<https://doi.org/10.17323/jil.2024.23990>

References / Список источников:

- Butler J. (2020) The Corporate Keepers of International Law. *American Journal of International Law*, vol. 114, no. 2, pp. 189–220.
- Crawford K. (2021) *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, New Heaven: Yale University Press.
- Donahoe E., Metzger M. (2019) Artificial Intelligence and Human Rights. *Journal of Democracy*, vol. 30, no. 2, pp. 115–126.
- Eubanks V. (2018) *Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor*, New York: St. Martin's Press.
- Filgueiras F. (2022) Artificial Intelligence Policy Regimes: Comparing Politics and Policy to National Strategies for Artificial Intelligence. *Global Perspectives*, vol. 3, no. 1, pp. 323–362. Available at: <https://doi.org/10.1525/gp.2022.32362>.
- Ho C. W-L., Caals K. (2024) How the EU AI Act Seeks to Establish an Epistemic Environment of Trust. *Asian Bioethics Review*, no. 16, pp. 345–372.
- Janig P. (2022) Extraterritorial Application of Human Rights. In: Binder C., Nowak M., Hofbauer J., Jang P. (eds.) *Elgar Encyclopedia of Human Rights*, vol II, Edward Elgar Publishing.
- Johnson J. (2019) Artificial Intelligence and Future Warfare: Implications for International Security. *Defense & Security Analysis*, vol. 35, no. 2, pp. 147–169.
- Kolfschooten H. v., Shachar C. (2023) The Council of Europe's AI Convention (2023–2024): Promises and pitfalls for health protection. *Health Policy*, vol. 138, pp. 1–5.
- Koskeniemi M. (2006) *From Apology to Utopia: The Structure of International Legal Argument*. Cambridge, Cambridge University Press.
- Lee J. (2022) *Artificial Intelligence and International Law*, Springer.
- Lütz F. (2024) The AI Act, gender equality and non-discrimination: what role for the AI office?. *ERA Forum* 25, pp. 79–95.
- Milanovic M. (2011) *Extraterritorial Application of Human Rights Treaties: Law, Principles, and Policy*, Oxford University Press.
- Nemitz P. (2018) Constitutional Democracy and Technology in the Age of Artificial Intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, № 2133. Available at: URL: <https://doi.org/10.1098/rsta.2018.0089> (accessed: 08.09.2024).
- Novelli C., Casolari F., Rotolo A. et al. (2024) AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act. *Digital Society*. vol. 3, no. 13, pp. 1–29.
- Rittberger B., Schimmelfennig F. (2007) *The Constitutionalization of the European Union*. Routledge.
- Rusinova V., Korotkov S. (2021) Mandatory Corporate Human Rights Due Diligence Models: Shooting Blanks? *Russian Law Journal*, vol. 9, № 4, pp. 33–71.