

УДК 341

DOI: 10.34076/207137972024_2_38

Е. Н. Мельникова¹

**НАДНАЦИОНАЛЬНОЕ ПРАВОВОЕ РЕГУЛИРОВАНИЕ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА КАК ИНСТРУМЕНТ
РАСПРЕДЕЛЕНИЯ ОТВЕТСТВЕННОСТИ ЗА ПРИЧИНЕННЫЙ ВРЕД
(ОПЫТ ЕВРОСОЮЗА)**

В рамках «цифрового» пакета документов, направленных на создание правового режима безопасного использования искусственного интеллекта (далее – ИИ), Евросоюзом предприняты попытки комплексного правового регулирования ИИ по двум направлениям: 1) установление административно-правовых обязанностей участников жизненного цикла систем и приложений ИИ; 2) создание гражданско-правового механизма распределения ответственности между участниками посредством установления соответствующих презумпций. В перечне обязанностей основных участников жизненного цикла систем и приложений ИИ, закрепленных в ключевом нормативном правовом документе (Регламенте ЕС об ИИ), четко прослеживается приверженность концепции значимого контроля за использованием ИИ. Концепция значимого контроля в совокупности с гражданско-правовыми инструментами, предложенными в проектах директив об ответственности за использование ИИ и об ответственности за дефектную продукцию, в идеале должны устранить «разрыв ответственности» за вред, причиненный ИИ, чего, однако, не удастся достичь ввиду ряда законодательных недоработок. В статье на примере ЕС проанализирован механизм наднационального правового

¹ Мельникова Елена Николаевна – соискатель Санкт-Петербургского государственного университета, юрисконсульт отдела юридического сопровождения деятельности в сфере экономики и финансов Юридического управления Национального исследовательского университета ИТМО (Санкт-Петербург), ORCID: 0000-0001-7263-5281, e-mail: melnikova_elenas@mail.ru.

регулирования ИИ как инструмент поиска сбалансированного распределения ответственности за вред, причиненный ИИ.

Ключевые слова: искусственный интеллект, «разрыв ответственности», значимый человеческий контроль за использованием ИИ, ответственность за дефектную продукцию в Европейском союзе, наднациональный механизм правового регулирования

Для цитирования: Мельникова Е. Н. Наднациональное правовое регулирование искусственного интеллекта как инструмент распределения ответственности за причиненный вред (опыт Евросоюза) // Российский юридический журнал. 2024. № 2. С. . DOI: 10.34076/20713797_2024_2_.

Введение

Стремительное создание и развитие технологий искусственного интеллекта (далее – ИИ) неизбежно, однако их использование может причинять вред жизни и здоровью человека и окружающей среде. Во всем мире предпринимаются попытки стимулирования ответственной разработки и использования технологий ИИ при помощи различных правовых средств. Первоочередной задачей представляется выработка механизма распределения ответственности за потенциальный вред, который может причинить ИИ, между лицами, принимающими участие в его создании и использовании: разработчиком, производителем, поставщиком, эксплуатантом (профессиональным пользователем) и др. В настоящей работе назовем этих лиц участниками жизненного цикла систем и приложений ИИ.

Использование технологий ИИ характеризуется опосредованием действий человека такими процессами, которые с учетом существующего уровня техники и человеческих способностей недостаточно прозрачны и объяснимы для поиска причинителя вреда. Это приводит в ряде случаев к «размыванию подотчетности» действий участников жизненного цикла системы или приложения ИИ и к такому нежелательному правовому явлению, как «разрыв ответственности».

При «разрыве ответственности» возложение ответственности за вред, причиненный использованием ИИ, на конкретное лицо (лиц) по правилам деликтной ответственности невозможно ввиду отсутствия таких элементов состава гражданско-правового нарушения, как причинно-следственная связь между вредом и действиями конкретных лиц, а также противоправность и виновность этих действий. Основным признаком, по которому можно судить о том, что при юридической квалификации юрист столкнется с «разрывом ответственности», – это затруднения в ходе технической экспертизы атрибутирования ошибки вывода результата системой ИИ, приведшей к неблагоприятным последствиям использования ИИ. Такие затруднения можно разбить на три вида:

1) имманентная ошибка модели ИИ – объективная невозможность атрибутирования ошибки конкретному лицу ввиду вероятностного характера вывода результата модели;

2) проблема «черного ящика» – случай, когда техническая экспертиза установила отсутствие имманентной ошибки модели ИИ, но не в состоянии атрибутировать неимманентную ошибку конкретному лицу;

3) проблема «многих рук» – случаи, когда работа системы ИИ, управляющей приложением ИИ, опосредуется действиями различных лиц, что ведет к субъективной невозможности атрибутирования ошибки.

Для устранения «разрыва ответственности» юридическими средствами требуется правовое регулирование ИИ, учитывающие особенности этих технологий.

Первые значимые попытки комплексного правового регулирования ИИ предприняты пока только в ЕС в рамках так называемого цифрового пакета документов. Один из его пяти блоков предусматривает правовое регулирование создания и использования ИИ посредством установления административно-правовых обязанностей участников жизненного цикла систем и приложений

ИИ в Регламенте ЕС об ИИ, принятом в марте 2024 (далее – Регламент)², и гражданско-правовых презумпций, закрепленных в проектах пересмотренной Директивы об ответственности за дефектную продукцию³ и Директиве об ответственности за использование ИИ⁴. Последняя обладает рядом недостатков, поэтому вряд ли будет принята в текущей редакции, однако представляет исследовательский интерес в связи с предложенными в ней презумпциями. Пересмотренная Директива об ответственности за дефектную продукцию с большой вероятностью будет принята в 2024 г.

Указанные документы представляют собой уникальную попытку устранения «разрыва ответственности» за вред, причиненный ИИ, посредством наднационального правового регулирования, однако они не лишены значимых недостатков.

Исследование европейского подхода к устранению «разрыва ответственности» данного вопроса является актуальным в связи с подготовкой модельного закона СНГ, направленного на правовое регулирование ИИ. Концепция этого закона 26 октября 2023 г. была одобрена Постоянной

² Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 July 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 (OJ L, 12.7.2024, p. 1–144).

³ Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) - Updated version in light of changes in the AI Act (COM/2022/496 final). URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52022PC0496> (дата обращения: 21.08.2024).

⁴ Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive). (2022/0303 (COD)). URL: https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf (дата обращения: 29.02.2024).

комиссией МПА СНГ по науке, образованию и культуре⁵, а сам модельный закон должен быть представлен к концу 2025 г. При подготовке этого документа желательно избежать недостатков, допущенных европейским законодателем.

Распределение ответственности за вред, причиненный ИИ, согласно Регламенту ЕС об ИИ

Регламент напрямую не регулирует вопросы распределения гражданско-правовой ответственности за неблагоприятные последствия использования ИИ: он устанавливает административно-правовой режим создания и использования ИИ посредством обозначения основных обязанностей участников жизненного цикла приложений ИИ, что, в свою очередь, влияет на формирование и развитие охранительных правоотношений, возникающих по поводу возмещения вреда, причиненного ИИ.

Регламент содержит ряд особенностей, касающихся предмета правового регулирования и круга лиц, на которых распространяется действие Регламента.

1. Особенности предмета правового регулирования. Регламент устанавливает правовые режимы деятельности по созданию и использованию: 1) систем ИИ (специального и общего назначения), 2) моделей ИИ общего назначения, 3) приложений ИИ.

Системы ИИ – это общее название ряда программных продуктов, основанных на технологиях ИИ. Из понятия системы ИИ, закрепленного в ст. 3 (1) Регламента, следует, что под действие Регламента подпадают только те системы ИИ, которые могут проявлять адаптивность после развертывания. Проявление адаптивности после развертывания означает, что в зависимости от входных данных, поступающих в систему в процессе ее использования, гиперпараметры моделей и правила принятия решений могут изменяться без

⁵ Одобрена концепция модельного закона, регулирующего технологии искусственного интеллекта // Межпарламентская Ассамблея СНГ. 2023. 23 окт. URL: https://iacis.ru/novosti/postoyannye_komissii/odobrena_kontseptciya_modelnogo_zakona_reguliruyushhego_tehnologii_iskusstvennogo_intellekta_1 (дата обращения: 12.06.2024).

участия человека⁶. Этот технический процесс – адаптация в процессе эксплуатации – обозначается также термином «самообучение»⁷. Соответственно, ключевым признаком системы ИИ, регулируемой Регламентом, является ее способность к самообучению.

2. *Особенности действия по кругу лиц.* Регламент, согласно ст. 2 (1) (с), действует экстерриториально. Он не регулирует права и обязанности потребителей, а вводит фигуру профессионального пользователя систем ИИ – эксплуатанта (ст. 3 (4) Регламента).

Отличительная особенность Регламента – распределение обязанностей в цепочке создания стоимости ИИ (ст. 28), где самой примечательной является фигура поставщика (ст. 3 (2)), которая как бы заслоняет собой разработчика систем и моделей ИИ. Поставщик – это не просто лицо, стоящее между наукой, производством и бизнесом и выполняющее роль «коммерциализатора» разработок. Как будет показано ниже, поставщик – это лицо, берущее на себя большую часть рисков использования ИИ в коммерческих целях: размещая на рынке систему ИИ или модель ИИ общего назначения под своим именем, поставщик, по сути, принимает на себя в том числе обязанности разработчика по соблюдению требований Регламента, поэтому разработчик не является субъектом регулирования этого акта. Любое лицо, вводящее в гражданский оборот систему или приложение ИИ под своим именем или вносящее в них существенные изменения или изменяющее предполагаемое назначение, считается поставщиком.

⁶ Мельникова Е. Н. Правовая квалификация модели машинного обучения на основе технического анализа этого процесса // Закон. 2023. № 6. С. 193–194. DOI: 10.37239/0869-4400-2023-20-6-187-196.

⁷ Krück C., Meßmer D. «AI Flash»: Regulation of artificial intelligence by the EU – What is an «AI system» under the EU AI Act? // SKW Schwarz. 2024. 29 Feb. URL: <https://www.lexology.com/library/detail.aspx?g=f8b59f2d-74f3-41f4-ad03-23f3d960c16b> (дата обращения: 05.03.2024).

Регламент классифицирует модели ИИ на модели общего назначения (это универсальный генеративный ИИ, например большие языковые модели) и модели ИИ специального назначения (все остальные «самообучающиеся» модели). Для моделей специального назначения особый правовой режим не предусмотрен, поскольку их правовое регулирование охватывается правовым режимом системы ИИ. Однако Регламент выделяет фигуру поставщика модели ИИ общего назначения (GPAIM), для деятельности которого ст. 52с Регламента установлен специальный правовой режим, где основной обязанностью поставщика GPAIM названа обязанность по ведению технической документации. GPAIM должна соответствовать ряду правил, при соблюдении которых поставщик не будет нести ответственность за ошибку GPAIM ввиду отсутствия возможности контроля за использованием модели. В данном случае отвечать будут либо поставщик, либо эксплуатант системы или приложения ИИ, в которые встроены GPAIM.

В настоящей работе рассмотрим только распределение обязанностей между поставщиком и эксплуатантом, которые являются *основными* участниками жизненного цикла систем и приложений ИИ.

Как представляется, важнейшей особенностью Регламента, представляющей интерес для исследования, является распределение обязанностей между поставщиком и эксплуатантом с позиции концепции значимого человеческого контроля за использованием ИИ. В российском праве эта концепция малоизвестна, поэтому требуется небольшое отступление для пояснений ее основных положений.

Концепция значимого человеческого контроля за использованием ИИ в международном праве

Концепция руководящего контроля привлекает внимание ученых давно и имеет межотраслевое значение. Так, уже с 60-х гг. XX в. эта концепция была

очень популярна в инженерии и психологии дорожного движения⁸ в контексте исследования взаимосвязи ответственности и контроля человека за собственными действиями, не опосредованными автономными процессами. Начиная с 80-х гг. XX в. взаимосвязь между возможностью осуществления контроля и моральной ответственностью активно исследовалась многими учеными⁹. Как представляется, наиболее четко обусловленность моральной ответственности возможностью осуществления контроля обозначена в компатибилистской теории «руководящего контроля» Дж. Фишера и М. Равиццы, обнародованной в 1998 г.¹⁰

Использование технологии ИИ значительно усложнило процесс установления подотчетности действий конкретных лиц и усилило эффект «размывания» подотчетности, который, конечно, возникает и без участия ИИ. Однако именно технологии ИИ способны затруднить атрибутирование ошибки конкретным лицам вплоть до полной невозможности установления лиц, которых можно привлечь к ответственности. Использование технологий ИИ в автономных летальных системах вооружения (далее – АЛСВ) делает проблему

⁸ *Michon J. A. A critical view of driver behavior models: what do we know, what should we do? // Human behavior and traffic safety / ed. by L. Evans, R. C. Schwing. N. Y.: Plenum Press, 1985. P. 500.*

⁹ *Bovens M., Miceli M. P. The quest for responsibility: Accountability and citizenship in complex organizations // Administrative Science Quarterly. Vol. 44. No. 4. P. 846–847. DOI: 10.2307/2667065. URL: https://www.researchgate.net/publication/275840305_The_Quest_for_Responsibility_Accountability_and_Citizenship_in_Complex_Organisations (дата обращения: 02.12.2023); *Collingridge D. The Social Control of Technology. L.: Frances Pinter, 1980. 200 p.; Danaher J. Robots, law and the retribution gap // Ethics and Information Technology. 2016. Vol. 18. P. 299–309; Pesch U. Engineers and Active Responsibility // Science and Engineering Ethics. 2015. Vol. 21. No. 4. P. 925–939. DOI: 10.1007/s11948-014-9571-7; Van de Poel I., Sand M. Varieties of responsibility: two problems of responsible innovation // Synthese. 2018. Vol. 198. No. 1. P. 4769–4787. DOI: 10.1007/s11229-018-01951-7.**

¹⁰ *Fischer J., Ravizza M. Responsibility and Control: A Theory of Moral Responsibility. Cambridge: Cambridge University Press, 1998. 277 p.*

«разрыва ответственности» особенно острой; «корень зла» видят в ослаблении контроля над АЛСВ¹¹, или «вакууме подотчетности»¹², который неустраним при текущем уровне техники. Это привлекло к изучению концепции значимого человеческого контроля за использованием ИИ юристов-международников и исследователей этики ИИ¹³. В последнее время данная концепция активно разрабатывается в международном праве ввиду потенциала ее положений для устранения «разрыва ответственности» юридическими средствами и актуальности для систем и приложений ИИ гражданского назначения¹⁴.

Создатели компатибилистской теории, однако, не определили условия для достижения морально значимого контроля человека над действиями, опосредованными использованием сложных автономных технологических систем. Иными словами, они еще четко не обозначили условия, при которых люди могут сохранять контроль и моральную ответственность за действия, опосредованные использованием роботизированных интеллектуальных систем

¹¹ *Matthias A.* The responsibility gap: Ascribing responsibility for the actions of learning automata // *Ethics and Information Technology*. 2004. Vol. 6. No. 3. P. 175. DOI: 10.1007/s10676-004-3422-1; *Sparrow R.* Killer robots // *Journal of Applied Philosophy*. 2007. Vol. 24. No. 1. P. 62–77. DOI: 10.1111/j.1468-5930.2007.00346.x; *Haag A., Hepkins F.* Mind the gap: The lack of accountability for killer robots // *Human Right Watch*. 2015. 9 Apr. URL: <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots> (дата обращения: 01.02.2024).

¹² *Heyns C.* Report of the Special Rapporteur on Extra-Judicial, Summary or Arbitrary Executions. Geneva: United Nations, 2014. 1 Apr. URL: <https://digitallibrary.un.org/record/771922?ln=ru> (дата обращения: 02.01.2024).

¹³ *Heyns C.* Op. cit.; *Meloni C.* State and individual responsibility for targeted killings by drones // *Drones and Responsibility: Legal, Philosophical and Socio-Technical Perspectives on Remotely Controlled Weapons* / ed. by E. Di Nucci, F. Santoni de Sio. N. Y.: Routledge, 2016. 228 p. DOI: 10.4324/9781315578187-3.

¹⁴ *Santoni de Sio F., van den Hoven J.* Meaningful Human Control over Autonomous Systems: A Philosophical Account // *Frontiers in Robotics and AI*. 2018. Vol. 5. No. 15. P. 11. DOI: 10.3389/frobt.2018.00015. URL: <https://www.frontiersin.org/journals/robotics-and-ai/articles/10.3389/frobt.2018.00015/full> (дата обращения: 02.01.2024).

с высоким уровнем автономии¹⁵. Значительный вклад в решение этой проблемы внес немецкий исследователь в области этики ИИ С. Робинсон. Он систематизировал наработки компатибилистов и выделил три условия значимого человеческого контроля¹⁶, при соблюдении которых моральная ответственность может быть вменена лицу в случае вреда, причиненного использованием ИИ.

Во-первых, контроль за ИИ может рассматриваться как значимый тогда, когда он осуществляется лицом, обладающим специальными знаниями, т. е. квалификацией, *потенциально* достаточной для предвидения последствий своих решений, опосредованных системой ИИ, что означает *осмысленный контроль*.

Во-вторых, тип контроля за использованием ИИ в контексте концепции значимого человеческого контроля должен быть именно *человеческим или человекоподобным*, что означает возможность реального контроля над системой или приложением ИИ с учетом скорости человеческой реакции. Назовем это свойство «человекоподобностью».

В-третьих, С. Робинсон различает контроль над входными данными и контроль над результатами или последствиями действий конкретной системы ИИ. Представляется, что контроль над входными данными должен охватываться понятием осмысленного контроля, т. е. относится к первому условию (осмысленный контроль). Что касается контроля над выходными данными, то Ф. Сантони де Сियो и Й. ван ден Ховен еще в 2018 г. определили необходимое условие *прослеживаемости*, согласно которому автономная система остается под значимым контролем человека тогда, когда спроектирована так, чтобы предоставлять возможность всегда отслеживать

¹⁵ *Santoni de Sio F., van den Hoven J.* Op. cit. P. 11.

¹⁶ *Robbins S.* The many meanings of meaningful human control // AI and Ethics. 2023. 24 July. P. 1–14. DOI: 10.1007/s43681-023-00320-6. URL: https://www.researchgate.net/publication/372583176_The_many_meanings_of_meaningful_human_control (дата обращения: 01.02.2024).

результаты ее операций по крайней мере одному человеку в цепочке проектирования и эксплуатации¹⁷.

Анализ норм международно-правовых актов¹⁸ и научных трудов философов и юристов¹⁹ позволяет прийти к выводу о двух вариантах толкования значимого человеческого контроля за использованием ИИ – узком и широком.

Если все три перечисленных условия значимого человеческого контроля соблюдаются (т. е. имеется квалифицированное лицо, осуществляющее осмысленный контроль за человекоразмерной системой, обеспечивающей прослеживаемость вывода результата), то речь идет о контроле в узком смысле («размывание подотчетности» отсутствует, соответственно, «разрыв ответственности» не образуется, так как причинитель вреда может быть установлен).

Если одно или несколько условий контроля не соблюдаются, то для устранения «разрыва ответственности» важно выработать понятие значимого человеческого контроля за использованием ИИ в его *широком значении*, когда

¹⁷ *Santoni de Sio F., van den Hoven J.* Op. cit. P. 1.

¹⁸ Проект рекомендации об этических аспектах искусственного интеллекта № SHS/BIO/ АНЕГ-AI/2020/4 REV.2: подготовлен 07.09.2020. URL: https://unesdoc.unesco.org/ark:/48223/pf0000373434_rus.locale=ru (дата обращения: 01.02.2024); Defence Artificial Intelligence Strategy. June 2022. NGO Article 36. URL: <https://article36.org/updates/new-uk-government-position-on-autonomous-weapons-recognises-that-lines-need-to-be-drawn-but-lacks-detail-or-signs-of-real-leadership/> (дата обращения: 02.01.2024); *Haag A., Hepkins F.* Op. cit.; *Schmitt M. N.* Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations. 2nd ed. Cambridge: Cambridge University Press, 2017. 598 p.

¹⁹ *Trabucco L.* What is meaningful human control, anyway? Cracking the code on autonomous weapons and human judgment // Modern War Institute. 2023. 21 Sept. URL: <https://mwi.westpoint.edu/what-is-meaningful-human-control-anyway-cracking-the-code-on-autonomous-weapons-and-human-judgment/> (дата обращения: 02.01.2024); *Bovens M., Miceli M. P.* Op. cit.; *Collingridge D.* Op. cit.; *Danaher J.* Op. cit.; *Pesch U.* Op. cit.; *Van de Poel I., Sand M.* Op. cit.

контролирующее лицо должно отвечать за чужие действия, в том числе за ошибки вывода результата системой ИИ, что противоречит основаниям деликтной ответственности. Вместе с тем входящее в практику опосредование принятия человеческих решений зачастую непредсказуемыми технологиями ИИ требует конструирования ответственности за выбор, что как нельзя лучше иллюстрирует высказывание: «Если вы сознательно решаете важный вопрос путем подбрасывания монеты, вы осознаете, что решили положиться на то, что, как вам прекрасно известно, является системой принятия случайных решений, и вы должны признать, что другие будут считать вас ответственным за результат выбора (и вы не можете их винить в этом)»²⁰.

Важнейшим положением концепции значимого человеческого контроля за использованием ИИ является принцип, согласно которому моральную ответственность за причиненный ИИ вред несет лицо, которое этот контроль осуществляет. Трансформация моральной ответственности в юридическую возможна только тогда, когда будет нормативно закреплена обязанность того или иного лица осуществлять значимый человеческий контроль за использованием системы или приложения ИИ, будут легально обозначены условия такого контроля, а также создан юридический механизм, позволяющий привлечь к ответственности лицо, осуществляющее контроль (т.е. разработаны юридические инструменты, направленные на устранение гражданско-правового «разрыва ответственности» за причиненный вред (ущерб)).

Легальное закрепление условий значимого человеческого контроля за использованием ИИ в Регламенте ЕС об ИИ

В Регламенте установлен перечень обязанностей поставщика и эксплуатанта системы или приложения ИИ, которые напрямую влияют на распределение ответственности за вред, причиненный их использованием.

1. Обязанности поставщика по обеспечению человекоподобности системы ИИ. Системы ИИ должны проектироваться и разрабатываться таким образом, чтобы физические лица могли эффективно контролировать их в

²⁰ *Santoni de Sio F., van den Hoven J. Op. cit. P. 5.*

течение периода использования (ст. 14 (1) Регламента). С этой целью поставщик системы должен определить соответствующие *меры по надзору за персоналом* до ее вывода на рынок или ввода в эксплуатацию. Такие меры должны гарантировать, что в системе есть *встроенные эксплуатационные ограничения*, которые не могут быть ею преодолены, и что система реагирует на действия человека-оператора, при том что физические лица, которым поручен человеческий надзор, обладают необходимой компетенцией, знаниями и полномочиями для выполнения этой роли (п. 48 преамбулы Регламента).

Поставщик обязан разрабатывать *инструкции по использованию системы ИИ* (ст. 13 (2) Регламента) в соответствии с ее предполагаемым *назначением или в условиях разумно предсказуемого неправильного использования* (п. 42а преамбулы Регламента). Инструкция должна предусматривать в том числе степень точности вывода результата системой ИИ (ст. 13 (2) (ii) Регламента). Последнее является инструментом снятия с поставщика ответственности в случае имманентной ошибки модели. Полнота документации и инструкций играет важнейшую роль в случае причинения вреда системой или приложением ИИ.

2. *Обязанности поставщика по обеспечению прослеживаемости системы ИИ.* В Регламенте требования, относящиеся к прослеживаемости (*traceability*), перечислены в ст. 12. Для обеспечения прослеживаемости поставщик обязан *вести журналы*, автоматически генерируемые системами ИИ тогда, когда система ИИ находится под его контролем (ст. 16 (d), ст. 14 Регламента). Поставщик обязан *вести документацию* (ст. 16 (c), ст. 18 Регламента), что обеспечивает прослеживаемость принятия решений системой ИИ и является важным источником сведений для технической экспертизы по выяснению причин причинения вреда системой ИИ. Согласно ст. 13 (1) Регламента информация, предоставляемая поставщиком эксплуатанту, должна гарантировать достаточную прозрачность функционирования систем ИИ, позволяющую эксплуатанту интерпретировать выходные данные системы и использовать их надлежащим образом.

Основная обязанность эксплуатанта – осуществление контроля над системой ИИ, который считается надлежащим при одновременном соблюдении двух условий, закрепленных в виде обязанностей в ст. 14 и 29 Регламента.

Согласно ст. 29 (1) Регламента эксплуатанты систем ИИ должны принимать соответствующие технические и организационные меры для обеспечения использования систем в соответствии с инструкциями по использованию.

Согласно ст. 29 (1) (a) Регламента лица, развертывающие систему ИИ, при осуществлении *контроля* над ней должны обеспечить, чтобы физические лица, назначенные для *человеческого надзора* за системами ИИ высокого риска, обладали необходимой компетенцией, знаниями и полномочиями, а также получали необходимую техническую поддержку. Следует подчеркнуть, что меры по человеческому надзору должен указать поставщик, однако приниматься они должны согласно закону и на *усмотрение* эксплуатанта (ст. 29 (2), (4) Регламента); последнее несколько ослабляет бремя ответственности поставщика.

Кроме того, эксплуатант обязан осуществлять контроль над входными данными (ст. 29 (2) Регламента). Соответственно, лица, осуществляющие человеческий надзор, должны быть достаточно компетентны для отбора входящих данных или эксплуатации системы в среде, соответствующей ее назначению.

Регламент разделяет понятия «контроль» и «человеческий надзор». Думается, что они соотносятся как общее и частное: человеческий надзор является одним из средств значимого человеческого контроля за использованием ИИ. В контексте Регламента контроль осуществляет эксплуатант, а человеческий надзор – назначенные им физические лица.

Как представляется, к осмысленному контролю можно отнести и обязанность эксплуатанта по осуществлению *оценки воздействия на основные права человека* (ст. 29a Регламента); при этом эксплуатанты могут

воспользоваться существующими оценками поставщиков. Это правило возлагает обязанность по оценке рисков (а следовательно, и ответственность за инциденты) обратно на поставщика.

Если система ИИ спроектирована так, что она поддается осуществлению человеческого надзора (т. е. является человекоразмерной) и обладает свойством прослеживаемости, то можно говорить о том, что технические пределы делегирования человеческих полномочий машине соблюдены.

Меры по соблюдению условий прослеживаемости и человекоразмерности системы ИИ входят в круг *обязанностей поставщика*, а меры по обеспечению осмысленного контроля относятся к *обязанностям эксплуатанта*. Значимый человеческий контроль за использованием ИИ осуществляется совместно двумя субъектами – поставщиком и эксплуатантом.

Распределение обязанностей по контролю за использованием системы ИИ носит ассиметричный характер: Регламент налагает на поставщика массу позитивных обязанностей, тогда как практически единственной обязанностью эксплуатанта является обязанность предпринять действия в случае материализации рисков, указанных поставщиком (ст. 29а (1) (f) Регламента); при этом свойство человекоразмерности системы ИИ должно позволить эти меры осуществить.

Юридические инструменты директив ЕС, направленные на устранение гражданско-правового «разрыва ответственности» за вред (ущерб), причиненный ИИ

Обязанности поставщика и эксплуатанта систем ИИ, установленные Регламентом, легли в основу гражданско-правовых инструментов, направленных на распределение ответственности за вред (ущерб), причиненный системой ИИ. Ряд этих инструментов закреплён в проекте Директивы об ответственности за использование ИИ. Определения, приведенные в ст. 2 Директивы, соответствуют определениям, закреплённым в ст. 3 Регламента.

Соблюдение участниками жизненного цикла систем ИИ обязанностей, установленных Регламентом, расценивается с точки зрения возможности применения презумпций, предусмотренных ст. 4 Директивой, как проявление осторожности. Несоблюдение поставщиком или эксплуатантом системы ИИ обязанностей, установленных Регламентом, расценивается как проявление неосторожности, что позволяет применить в отношении нарушителя – потенциального причинителя вреда соответствующие презумпции (о них ниже). Доказанное несоблюдение обязанностей, установленных Регламентом, должно привести к выявлению нарушителя, в этом случае «разрыв ответственности» не образуется. Сложности возникают тогда, когда все обязанности, установленные Регламентом и иными законами, соблюдены, а вред, тем не менее, причинен. Тогда при невозможности атрибутирования ошибки посредством экспертизы возникает «разрыв ответственности».

Как отмечает С. В. Никитенко, «если автономная система совершила действие на основании собственного решения, то стандартных правил об объективной или безвиновной ответственности недостаточно, поскольку традиционные институты ответственности не учитывают случаи нанесения ущерба роботами, способными к самообучению. Данная позиция также поддерживается многими авторами²¹, что указывает на необходимость адаптации деликтной ответственности к ИИ-технологиям»²².

В ЕС разработали презумпции, представляющие собой «полумеры», направленные на устранение «разрыва ответственности».

Директива об ответственности за использование ИИ ставит задачей выявление причинителя вреда и направлена на облегчение бремени доказывания объективной стороны гражданского правонарушения, а именно

²¹ Цит. по: *Никитенко С. В.* Международно-правовое регулирование использования искусственного интеллекта в области медицины: дис. ... канд. юрид. наук. СПб., 2023. С. 116; *Толочко О. Н.* Тенденции правового регулирования объектов и технологий, связанных с искусственным интеллектом // Юстиция Беларуси. 2019. № 3. С. 35–39.

²² *Никитенко С. В.* Указ. соч. С. 116.

наличия причинно-следственной связи между применением ИИ и неблагоприятными последствиями²³.

Директива в ст. 4 вводит опровержимую презумпцию причинно-следственной связи между действиями ответчика и ошибкой, повлекшей причинение вреда. Презумпция применяется судом при соблюдении ряда условий, ключевым из которых является необходимость доказывания вины ответчика. Такая вина может быть установлена, например, в случае несоблюдения ответчиком обязанностей, предусмотренных Регламентом или другими правилами ЕС, что приравнивается к нарушению обязанности проявлять *осторожность* (ст. 4 (1) (а) Директивы).

Поставщик может защититься от иска, предъявив доказательства соблюдения условий человекоразмерности и прослеживаемости системы ИИ, создания системы управления рисками, соответствующей требованиям гл. 2 Регламента, соблюдения иных обязанностей, закрепленных Регламентом. Дополнительным аргументом могут служить результаты постмаркетингового мониторинга, демонстрирующие отсутствие серьезных инцидентов. Тем самым поставщик системы ИИ может доказать надлежащее исполнение обязанности проявлять осторожность и освободиться от ответственности.

Эксплуатант может сослаться на надлежащее выполнение инструкции поставщика, соблюдение обязанностей по мониторингу и использованию системы ИИ по назначению.

Поставщик модели ИИ общего назначения вообще не входит в число ответчиков, а если все же будет привлечен к делу, то с большой вероятностью укажет на то, что ошибка вызвана действиями поставщика или эксплуатанта системы ИИ. Например, в делах, связанных с ДТП с участием высокоавтоматизированных транспортных средств, судам приходится

²³ Vasudevan A. Who Is Liable for AI-Driven Accidents? The Law Is Still Emerging to establish negligence, a plaintiff needs to prove causation // CIGI. 2023. 21 June. URL: <https://www.cigionline.org/articles/who-is-liable-for-ai-driven-accidents-the-law-is-still-emerging/> (дата обращения: 17.08.2023).

устанавливать, следует ли привлекать к ответственности поставщика за некачественную разработку или причиной аварии стала невнимательность водителя во время вождения²⁴. В литературе можно встретить дискуссию о том, нужно ли привлекать врачей к ответственности за следование неточным или вводящим в заблуждение рекомендациям систем ИИ при диагностике и лечении пациентов²⁵. В этих делах поставщик системы ИИ к ответственности не привлекался.

Директива об ответственности за использование ИИ не отвечает на вопрос, кто будет нести ответственность за вред (ущерб) в случае надлежащего исполнения *всех* наложенных Регламентом обязательств и поставщиком, и эксплуатантом. На практике, вероятно, вопрос всегда будет разрешаться в чью-то пользу, поскольку вряд ли ответчикам удастся представить систему безупречного документирования. Однако при доказанном соблюдении всех обязанностей ответчиком «разрыв ответственности» не устраняется.

Директива не исключает «разрыв ответственности», как представляется, по двум причинам: а) она не помогает определить фигуру ответчика тогда, когда техническая экспертиза не позволяет достоверно выявить причину и «автора» ошибки и при этом все законодательные требования ответчиком соблюдены; б) презумпция, установленная Директивой, достаточно легко опровергается.

²⁴ *Suber S. H., Saxon M.* First Lawsuit Filed for Tesla Autopilot-Related Death Involving a Pedestrian // Lexology. 2020. 16 June. URL: <https://www.lexology.com/library/detail.aspx?g=c6019770-6bb2-4680-b7ad-24d964637640> (дата обращения: 17.08.2023).

²⁵ *Smith H., Fotheringham K.* Artificial intelligence in clinical decision-making: Rethinking liability // Medical Law International. 2020. Vol. 20. Is. 2. P. 131–154. DOI: 10.1177/0968533220945766. URL: https://research-information.bris.ac.uk/ws/portalfiles/portal/248250139/AI_and_clinical_decision_making_June_2020_rev3.pdf (дата обращения: 17.08.2023).

С 1985 г. в Евросоюзе действует Директива ЕС об ответственности за дефектную продукцию²⁶, предусматривающая *ответственность производителя за ущерб, причиненный потребителям в результате дефектов продукции независимо от того, появился ли дефект по вине производителя*. Согласно ст. 4 (1) данной директивы истец обязан доказать: 1) понесенный ущерб; 2) дефектность продукта; 3) причинно-следственную связь между дефектом и ущербом. Последние два элемента состава в случае с ИИ установить крайне сложно²⁷, поэтому 24 января 2024 г. Советом ЕС была согласована новая редакция директивы (далее – пересмотренная Директива).

В ст. 9 (3) и (4) пересмотренной Директивы содержится опровержимая презумпция, согласно которой если вред причинен продуктом (системой или приложением ИИ), то *наличие дефекта изделия и причинно-следственная связь между дефектом и вредом (ущербом) предполагаются*. Данная презумпция легко опровержима. Для этого, исходя из ст. 9 (2) пересмотренной Директивы, ответчик должен доказать соответствие продукта обязательным требованиям и неочевидность неисправности изделия при нормальном использовании или при обычных обстоятельствах.

Примечательно, что пересмотренная Директива стремится к возложению ответственности на лиц, осуществляющих контроль за использованием продукта. В ней применена концепция руководящего контроля, что следует из закрепленного в ст. 6 перечня обстоятельств, используемых для оценки дефектности продукта: а) *инструкции по установке, использованию и техническому обслуживанию*; б) *разумно предсказуемое использование*

²⁶ Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products. (Product Liability Directive). URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31985L0374> (дата обращения: 29.02.2024).

²⁷ *Buiten M., de Streel A., Peitz M.* The Law and Economics of AI Liability // Computer Law and Security Review. 2023. Vol. 48. Article 105794. DOI: 10.1016/j.clsr.2023.105794. URL: <https://www.sciencedirect.com/science/article/pii/S0267364923000055?via%3Dihub> (дата обращения: 29.02.2024).

продукта; с) влияние на продукт способности *продолжать обучение и приобретать новые признаки* после выхода на рынок; d) влияние на продукт *других продуктов*, которые, как можно обоснованно ожидать, будут использоваться вместе с данным продуктом; e) *сохранение производителем контроля над продуктом* после его введения в гражданский оборот; f) время, *когда продукт вышел из-под контроля производителя* (ст. 6 (1) пересмотренной Директивы). Тем самым изначальный производитель системы или приложения ИИ выводится из круга потенциальных ответчиков и надлежащим ответчиком становится лицо, осуществляющее контроль над продуктом.

Соответственно, презумпция наличия дефекта изделия и причинно-следственной связи между дефектом и вредом (ущербом) позволяет *устранить «разрыв ответственности»* тогда, когда *производитель или экономический оператор не справились с опровержением этой презумпции*. Так, чтобы освободиться от ответственности, производитель, экономический оператор должны доказать, что они не сохраняли контроль над продуктом тогда, когда проявился дефект. В этом случае они будут признаны ненадлежащими ответчиками, после чего истец может предъявить требование к *контролирующему лицу*, в частности к тому, кто произвел в отношении продукта действия, перечисленные в ст. 6 пересмотренной Директивы.

Экономические операторы, ответственные за дефектную продукцию, названы в ст. 7 пересмотренной Директивы. Среди них нет лиц, которые произвели изменения продукта, поэтому в отношении таких лиц не будет действовать презумпция, закрепленная в ст. 4 пересмотренной Директивы. И тогда истец столкнется со сложностями в установлении причинно-следственной связи между применением ИИ и причинением вреда согласно Директиве об ответственности за использование ИИ. В случае опровержения этой презумпции ответчиком – контролирующим лицом при наличии у последнего доказательств надлежащего исполнения им всех обязательных требований и инструкций поставщика образуется «разрыв ответственности», который

невозможно устранить юридическими инструментами, предложенными в обеих директивах.

Наднациональное регулирование распределения ответственности за вред, причиненный приложениями ИИ

Анализ положений Регламента и обеих директив ЕС показал, что содержащиеся в них инструменты не в силах полностью устранить «разрыв ответственности». Несмотря на убедительные доводы, приводимые сторонниками подхода, согласно которому снятие обязанности по контролю с пользователей (которые ожидают, что автономные системы будут выполнять свою работу независимо) должно повлечь передачу ответственности производителям²⁸, можно говорить о бесперспективности возложения ответственности на производителя, который легко может снять ее с себя, оставив «разрыв ответственности» незакрытым. Пересмотренная Директива лишь намечает путь возложения ответственности на контролирующее лицо, но не доводит его до конца, т. е. до полного устранения «разрыва ответственности». Пересмотренная Директива (как и изначальный документ 1985 г.) не применяется к хозяйствующим субъектам (в том числе к поставщикам и эксплуатантам в терминологии Регламента), а регулирует правоотношения экономических операторов с физическими лицами, в том числе с потребителями, поэтому «разрыв ответственности» в правоотношениях между юридическими лицами остается незакрытым.

Ценным опытом ЕС в правовом регулировании ИИ является распределение в Регламенте обязанностей между поставщиком и эксплуатантом с учетом положений концепции значимого человеческого контроля. Через механизм применения судом презумпций ответственность может «перемещаться» от поставщика к эксплуатанту и обратно – в зависимости от того, кто лучше справляется с бременем доказывания. Если оба представили идеальные доказательства проявления осторожности и при этом

²⁸ *Buiten M.* Product Liability for Defective AI // *European Journal of Law and Economics*. 2024. Vol. 57. P. 268.

экспертиза не в состоянии атрибутировать ошибку, то возникает «разрыв ответственности», который невозможно устранить предложенными в ЕС правовыми средствами. Такая ситуация нежелательна, выявленное «узкое место» в законодательстве ЕС должно быть учтено законодателями других региональных союзов в ходе формирования собственного правового регулирования ИИ. Вместо того чтобы при «разрыве ответственности» возлагать ответственность на контролирующее лицо, европейский законодатель предпочел оставить разрешение данного вопроса на усмотрение суда. Негативный опыт ЕС в части презумпций не должен воспроизводиться в других правовых системах.

Возложение ответственности на лицо, осуществляющее значимый человеческий контроль за использованием ИИ, требует нового юридического подхода к установлению состава деликта с позиции концепции ответственности за выбор средств принятия решений.

Очевидно, что внедрение в законодательство ряда стран новаторского правила, особенно такого, которое может не в полной мере разделяться отдельными группами влияния внутри некоторых государств, в короткое время возможно только «сверху вниз», т. е. посредством механизма международно-правового регулирования.

Мы лишь наметили один из возможных правовых способов устранения «разрыва ответственности» для того, чтобы проанализировать пути формирования нормативно-правового регулирования распределения ответственности за вред, причиненный приложениями ИИ. Важно не только разработать вариант решения правовой проблемы с точки зрения теории и сформулировать «правильные» правоположения. Не менее значимо, чтобы эти правоположения были приемлемы для большинства на целевой территории.

На примере комплексного правового регулирования ИИ в ЕС становится очевидным значение наднационального механизма в процессе поиска компромиссного варианта распределения ответственности за вред, причиненный ИИ. Динамику изменений проекта нормативного акта в ходе

законодательного процесса легко проследить (по сравнению с анализом неидентичного содержания национальных правовых актов, принятых в разное время), что и позволяет любому заинтересованному лицу понять расстановку сил и оценить правовые перспективы того или иного управленческого решения. Результатом достигнутого в ЕС компромисса в распределении ответственности за использование ИИ стало сосредоточение обязанностей и ответственности у поставщика систем ИИ, что стало причиной «узких мест» в правовом регулировании ИИ и существенным препятствием к устранению «разрыва ответственности» за вред, причиненный ИИ.

Следует в очередной раз подчеркнуть, что в сфере регулирования ИИ фрагментация правового регулирования и его дифференциация по странам в рамках интеграционного образования недопустима. Для трансграничного нематериального феномена нужны трансграничные средства правового регулирования. Только на наднациональном уровне можно ввести в юридический обиход новый подход к решению правовой проблемы, комплексно урегулировав правоотношения в сфере ИИ.

Выводы

1. Регламент содержит административно-правовые обязанности участников жизненного цикла систем ИИ, определяющие распределение гражданско-правовой ответственности за вред, причиненный ИИ. Характер этих обязанностей позволяет сделать вывод, что европейский законодатель формировал положения Регламента с учетом концепции значимого человеческого контроля за использованием ИИ. Вместе с тем обязанности по осуществлению такого контроля разделены между поставщиком систем ИИ и эксплуатантом систем ИИ так, что основная их часть оказывается возложенной на поставщика, что приводит к концентрации ответственности на поставщике системы или приложения ИИ – лице, которое не способно в полной мере контролировать их использование.

2. Гражданско-правовые презумпции, установленные Директивой об ответственности за использование ИИ и Директивой об ответственности

производителя за дефектную продукцию, в целом облегчают бремя доказывания наличия дефекта и причинно-следственной связи между дефектом и использованием ИИ, но не способны полностью устранить гражданско-правовой «разрыв ответственности».

3. Анализ проектов документов ЕС, направленных на регулирование ИИ, позволяет прийти к выводу о сохранении «разрыва ответственности» за вред (ущерб), причиненный ИИ. Вместе с тем строго детерминированное влияние административно-правовых обязанности участников жизненного цикла систем ИИ на распределение гражданско-правовой ответственности, четко прослеживаемое в законодательстве ЕС, может послужить ценным опытом для формирования нормативно-правового регулирования в других интеграционных образованиях, например в ЕАЭС.

4. Возложение ответственности на лицо, осуществляющее значимый человеческий контроль за использованием ИИ, представляется одним из возможных вариантов сбалансированного распределения ответственности за причиненный вред. Этот подход заслуживает внимания при формировании правовой базы для регулирования ИИ как на уровне ЕАЭС, так и на уровне отдельных государств-членов.

5. Как показывает опыт ЕС, введение на территории группы государств юридического инструмента, способного скорректировать применительно к ИИ классический принцип ответственности производителя за свою продукцию посредством смещения ответственности в сторону контролирующего профессионального пользователя, может быть эффективно осуществлено с помощью наднационального регулирования. Именно такой механизм обеспечивает наилучшее качество законодательных решений и позволяет избежать фрагментации законодательства и правовой неопределенности, недопустимой в данной сфере общественных отношений.

6. Опыт ЕС демонстрирует растущее понимание необходимости регулирования ИИ прежде всего на наднациональном уровне. Это актуально и для России, и для ЕАЭС, где активно развивается использование ИИ.

Список литературы

Мельникова Е. Н. Правовая квалификация модели машинного обучения на основе технического анализа этого процесса // Закон. 2023. № 6. С. 187–196. DOI: 10.37239/0869-4400-2023-20-6-187-196.

Никитенко С. В. Международно-правовое регулирование использования искусственного интеллекта в области медицины: дис. ... канд. юрид. наук. СПб., 2023. 223 с.

Толочко О. Н. Тенденции правового регулирования объектов и технологий, связанных с искусственным интеллектом // Юстиция Беларуси. 2019. № 3. С. 35–39.

Bovens M., Miceli M. P. The quest for responsibility: Accountability and citizenship in complex organizations // Administrative Science Quarterly. Vol. 44. No. 4. P. 846–847. DOI: 10.2307/2667065. URL: https://www.researchgate.net/publication/275840305_The_Quest_for_Responsibility_Accountability_and_Citizenship_in_Complex_Organisations (дата обращения: 02.12.2023).

Buiten M. Product Liability for Defective AI // European Journal of Law and Economics. 2024. Vol. 57. P. 239–273.

Buiten M., de Streef A., Peitz M. The Law and Economics of AI Liability // Computer Law and Security Review. 2023. Vol. 48. Article 105794. DOI: 10.1016/j.clsr.2023.105794. URL: <https://www.sciencedirect.com/science/article/pii/S0267364923000055?via%3Dihub> (дата обращения: 29.02.2024).

Collingridge D. The Social Control of Technology. L.: Frances Pinter, 1980. 200 p.

Danaher J. Robots, law and the retribution gap // Ethics and Information Technology. 2016. Vol. 18. P. 299–309.

Fischer J., Ravizza M. Responsibility and Control: A Theory of Moral Responsibility. Cambridge: Cambridge University Press, 1998. 277 p.

Haag A., Hepkins F. Mind the gap: The lack of accountability for killer robots // Human Right Watch. 2015. 9 Apr. URL: <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots> (дата обращения: 01.02.2024).

Heyns C. Report of the Special Rapporteur on Extra-Judicial, Summary or Arbitrary Executions. Geneva: United Nations, 2014. 1 Apr. URL: <https://digitallibrary.un.org/record/771922?ln=ru> (дата обращения: 02.01.2024).

Krück C., Meßmer D. «AI Flash»: Regulation of artificial intelligence by the EU – What is an «AI system» under the EU AI Act? // SKW Schwarz. 2024. 29 Feb. URL: <https://www.lexology.com/library/detail.aspx?g=f8b59f2d-74f3-41f4-ad03-23f3d960c16b> (дата обращения: 05.03.2024).

Matthias A. The responsibility gap: Ascribing responsibility for the actions of learning automata // Ethics and Information Technology. 2004. Vol. 6. No. 3. P. 175–183. DOI: 10.1007/s10676-004-3422-1.

Meloni C. State and individual responsibility for targeted killings by drones // Drones and Responsibility: Legal, Philosophical and Socio-Technical Perspectives on Re-motely Controlled Weapons / ed. by E. Di Nucci, F. Santoni de Sio. N. Y.: Routledge, 2016. 228 p. DOI: 10.4324/9781315578187-3.

Michon J. A. A critical view of driver behavior models: what do we know, what should we do? // Human behavior and traffic safety / ed. by L. Evans, R. C. Schwing. N. Y.: Plenum Press, 1985. P. 485–524.

Pesch U. Engineers and Active Responsibility // Science and Engineering Ethics, 2015. Vol. 21. No. 4. P. 925–939. DOI: 10.1007/s11948-014-9571-7.

Robbins S. The many meanings of meaningful human control // AI and Ethics. 2023. 24 July. P. 1–14. DOI: 10.1007/s43681-023-00320-6. URL: https://www.researchgate.net/publication/372583176_The_many_meanings_of_meaningful_human_control (дата обращения: 01.02.2024).

Santoni de Sio F., van den Hoven J. Meaningful Human Control over Autonomous Systems: A Philosophical Account // Frontiers in Robotics and AI. 2018. Vol. 5. No. 15. P. 1–14. DOI: 10.3389/frobt.2018.00015. URL:

<https://www.frontiersin.org/journals/robotics-and-ai/articles/10.3389/frobt.2018.00015/full> (дата обращения: 02.01.2024).

Schmitt M. N. Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations. 2nd ed. Cambridge: Cambridge University Press, 2017. 598 p. URL: <https://www.cambridge.org/core/books/tallinn-manual-20-on-the-international-law-applicable-to-cyber-operations/E4FFD83EA790D7C4C3C28FC9CA2FB6C9> (дата обращения: 01.11.2023).

Smith H., Fotheringham K. Artificial intelligence in clinical decision-making: Rethinking liability // *Medical Law International*. 2020. Vol. 20. Is. 2. P. 131–154. DOI: 10.1177/0968533220945766. URL: https://research-information.bris.ac.uk/ws/portalfiles/portal/248250139/AI_and_clinical_decision_making_June_2020_rev3.pdf (дата обращения: 17.08.2023).

Sparrow R. Killer robots // *Journal of Applied Philosophy*. 2007. Vol. 24. No. 1. P. 62–77. DOI: 10.1111/j.1468-5930.2007.00346.x.

Suber S. H., Saxon M. First Lawsuit Filed for Tesla Autopilot-Related Death Involving a Pedestrian // *Lexology*. 2020. 16 June. URL: <https://www.lexology.com/library/detail.aspx?g=c6019770-6bb2-4680-b7ad-24d964637640> (дата обращения: 17.08.2023).

Trabucco L. What is meaningful human control, anyway? Cracking the code on autonomous weapons and human judgment // *Modern War Institute*. 2023. 21 Sept. URL: <https://mwi.westpoint.edu/what-is-meaningful-human-control-anyway-cracking-the-code-on-autonomous-weapons-and-human-judgment/> (дата обращения: 02.01.2024).

Van de Poel I., Sand M. Varieties of responsibility: two problems of responsible innovation // *Synthese*. 2018. Vol. 198. No. 1. P. 4769–4787. DOI: 10.1007/s11229-018-01951-7.

Vasudevan A. Who Is Liable for AI-Driven Accidents? The Law Is Still Emerging to establish negligence, a plaintiff needs to prove causation // *CIGI*. 2023. 21 June. URL: <https://www.cigionline.org/articles/who-is-liable-for-ai-driven-accidents-the-law-is-still-emerging/> (дата обращения: 17.08.2023).

UDC 341

DOI: 10.34076/20713797_2024_2_

SUPRANATIONAL LEGAL REGULATION OF ARTIFICIAL INTELLIGENCE AS A TOOL FOR DISTRIBUTING RESPONSIBILITY FOR HARM CAUSED (THE EXPERIENCE OF THE EUROPEAN UNION)

Melnikova Elena – applicant, Saint Petersburg State University (Saint Petersburg), legal counsel, National Research University ITMO, ORCID: 0000-0001-7263-5281, e-mail: melnikova_elen5@mail.ru.

Within the framework of the «Digital Package» aimed at creating a legal regime for the safe use of Artificial Intelligence (hereinafter – AI), the European Union has made significant attempts to comprehensively regulate AI in two directions: 1) the establishment of administrative and legal responsibilities of participants in the life cycle of AI systems and applications, and 2) a civil law mechanism for distributing responsibility between them by establishing appropriate presumptions. In the list of responsibilities of the main participants in the life cycle of AI systems and applications, enshrined in the key regulatory document (AI Act), the European legislator clearly follows the provisions of the concept of meaningful control over the use of AI, which, together with the civil law instruments proposed in the draft directive on responsibility for AI and the revised directive on responsibility for defective products, ideally, should eliminate the «responsibility gap» for damage caused by AI, which, however, It is not possible to achieve this due to a number of legislative flaws. In this article, using the example of the EU, the mechanism of supranational legal regulation of AI is analyzed as a tool for finding a balanced distribution of responsibility for harm caused by AI. The supranational mechanism of legal regulation of AI is justified as the optimal way to make legislative decisions in the field of AI.

Key words: artificial intelligence, responsibility gap, meaningful human control over the use of AI, AI Act, AI Liability Directive, liability for defective products in the European Union, supranational legal regulation mechanism

For citation: Melnikova E. (2024) Supranational legal regulation of artificial intelligence as a tool for distributing responsibility for harm caused (the experience of the European Union). In *Rossiiskii yuridicheskii zhurnal*, no. 2, pp. , DOI: 10.34076/20713797_2024_2_.

References

Bovens M., Miceli M. P. (1999) The quest for responsibility: Accountability and citizenship in complex organizations. In *Administrative Science Quarterly*, vol. 44, no. 4, pp. 846–847, DOI: 10.2307/2667065, available at: https://www.researchgate.net/publication/275840305_The_Quest_for_Responsibility_Accountability_and_Citizenship_in_Complex_Organisations (accessed: 02.12.2023).

Buiten M. (2024) Product Liability for Defective AI. In *European Journal of Law and Economics*, vol. 57, pp. 239–273.

Buiten M., de Streel A., Peitz M. (2023) The Law and Economics of AI Liability. In *Computer Law and Security Review*, vol. 48, article 105794, DOI: 10.1016/j.clsr.2023.105794, available at: <https://www.sciencedirect.com/science/article/pii/S0267364923000055?via%3Dihub> (accessed: 29.02.2024).

Collingridge D. (1980) *The Social Control of Technology*. London, Frances Pinter, 200 p.

Danaher J. (2016) Robots, law and the retribution gap. In *Ethics and Information Technology*, vol. 18, pp. 299–309.

Fischer J., Ravizza M. (1998) *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge, Cambridge University Press, 277 p.

Haag A., Hepkins F. (2015) Mind the gap: The lack of accountability for killer robots. In *Human Right Watch*. 9 Apr., available at:

<https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots>
(accessed: 01.02.2024).

Heyns C. (2014) *Report of the Special Rapporteur on Extra-Judicial, Summary or Arbitrary Executions*. Geneva, United Nations, 1 Apr., available at: <https://digitallibrary.un.org/record/771922?ln=ru> (accessed: 02.01.2024).

Krück C., Meßmer D. (2024) «AI Flash»: Regulation of artificial intelligence by the EU – What is an «AI system» under the EU AI Act? In *SKW Schwarz*, 29 Feb., available at: <https://www.lexology.com/library/detail.aspx?g=f8b59f2d-74f3-41f4-ad03-23f3d960c16b> (accessed: 05.03.2024).

Matthias A. (2004) The responsibility gap: Ascribing responsibility for the actions of learning automata. In *Ethics and Information Technology*, vol. 6, no. 3, pp. 175–183, DOI: 10.1007/s10676-004-3422-1.

Melnikova E. N. (2023) Pravovaya kvalifikaciya modeli mashinnogo obucheniya na osnove tehničeskogo analiza etogo processa [The legal nature of the machine learning model]. In *Zakon*, no. 6, pp. 187–196.

Meloni C. (2016) State and individual responsibility for targeted killings by drones. In Di Nucci E., Santoni de Sio F. (Eds.) *Drones and Responsibility: Legal, Philosophical and Socio-Technical Perspectives on Re-motely Controlled Weapons*. New York, Routledge, 228 p., DOI: 10.4324/9781315578187-3.

Michon J. A. (1985) A critical view of driver behavior models: what do we know, what should we do? In Evans L., Schwing R. C. (Eds.) *Human behavior and traffic safety*. New York, Plenum Press, pp. 485–524.

Nikitenko S. V. (2023) *Mezhdunarodno-pravovoe regulirovanie ispol'zovaniya iskusstvennogo intellekta v oblasti meditsiny: dis. ... kand. jurid. nauk* [International legal regulation of the use of artificial intelligence in the field of medicine: a candidate of legal sciences thesis]. Saint Peterburg, 223 p.

Pesch U. (2015) Engineers and Active Responsibility. In *Science and Engineering Ethics*, vol. 21, no. 4, pp. 925–939, DOI: 10.1007/s11948-014-9571-7.

Robbins S. (2023) The many meanings of meaningful human control. In *AI and Ethics*, 24 July, pp. 1–14, DOI: 10.1007/s43681-023-00320-6, available at:

https://www.researchgate.net/publication/372583176_The_many_meanings_of_meaningful_human_control (accessed: 01.02.2024).

Santoni de Sio F., van den Hoven J. (2010) Meaningful Human Control over Autonomous Systems: A Philosophical Account. In *Frontiers in Robotics and AI*, vol. 5, no. 15, pp. 1–14, DOI: 10.3389/frobt.2018.00015, available at: <https://www.frontiersin.org/journals/robotics-and-ai/articles/10.3389/frobt.2018.00015/full> (accessed: 02.01.2024).

Schmitt M. N. (2017) *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*, 2nd ed. Cambridge, Cambridge University Press, 598 p., available at: <https://www.cambridge.org/core/books/tallinn-manual-20-on-the-international-law-applicable-to-cyber-operations/E4FFD83EA790D7C4C3C28FC9CA2FB6C9> (accessed: 01.11.2023).

Smith H., Fotheringham K. (2020) Artificial intelligence in clinical decision-making: Rethinking liability. In *Medical Law International*, vol. 20, is. 2, pp. 131–154, DOI: 10.1177/0968533220945766, available at: https://research-information.bris.ac.uk/ws/portalfiles/portal/248250139/AI_and_clinical_decision_making_June_2020_rev3.pdf (accessed: 17.08.2023).

Sparrow R. (2007) Killer robots. In *Journal of Applied Philosophy*, vol. 24, no. 1, pp. 62–77, DOI: 10.1111/j.1468-5930.2007.00346.x.

Suber S. H., Saxon M. (2020) First Lawsuit Filed for Tesla Autopilot-Related Death Involving a Pedestrian. In *Lexology*, 16 June, available at: <https://www.lexology.com/library/detail.aspx?g=c6019770-6bb2-4680-b7ad-24d964637640> (accessed: 17.08.2023).

Tolochko O. N. (2019) Tendentsii pravovogo regulirovaniya ob"ektov i tekhnologii, svyazannykh s iskusstvennym intellektom [Trends in legal regulation of objects and technologies related to artificial intelligence]. In *Yustitsiya Belarusi*, no. 3, pp. 35–39.

Trabucco L. (2023) What is meaningful human control, anyway? Cracking the code on autonomous weapons and human judgment. In *Modern War Institute*. 21 Sept., available at: <https://mwi.westpoint.edu/what-is-meaningful-human-control->

anyway-cracking-the-code-on-autonomous-weapons-and-human-judgment/
(accessed: 02.01.2024).

Van de Poel I., Sand M. (2018) Varieties of responsibility: two problems of responsible innovation. In *Synthese*, vol. 198, no. 1, pp. 4769–4787, DOI: 10.1007/s11229-018-01951-7.

Vasudevan A. (2023) Who Is Liable for AI-Driven Accidents? The Law Is Still Emerging to establish negligence, a plaintiff needs to prove causation. In *CIGI*, 21 June, available at: <https://www.cigionline.org/articles/who-is-liable-for-ai-driven-accidents-the-law-is-still-emerging/> (accessed: 17.08.2023).