

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ПРАВОВОЙ СТАТУС ИЛИ ПРАВОВОЙ РЕЖИМ?

© 2022 г. Ю. М. Батури¹, *, С. В. Полубинская², **

¹Московский государственный университет им. М. В. Ломоносова

²Институт государства и права Российской академии наук, г. Москва

*E-mail: baturin@ihst.ru

**E-mail: svepol@yandex.ru

Поступила в редакцию 28.06.2022 г.

Аннотация. В России, как и в мире, искусственный интеллект, роботы и роботизированные технологии внедряются в различные области экономики, государственной и общественной жизни. Разнообразие сфер применения, в которых уже используются и будет использоваться искусственный интеллект в различных формах, оказывает воздействие на правовую систему и нуждается в разработке адекватного правового регулирования, учитывающего особенности этих принципиально новых и ранее неизвестных технологий. Решение такой сложной задачи, в свою очередь, требует теоретического осмысления ряда концептуальных вопросов, способных повлиять на успешность или неудачу правовой регламентации в рассматриваемой области общественных отношений.

В статье анализируются законы робототехники А. Азимова и этические нормы, предлагаемые для искусственного интеллекта в Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники на период до 2024 года. Рассматриваются вопросы уголовно-правовой оценки поведения роботов, предлагаются понимание и определение таких понятий, как «этика» роботов и «правомерность» их поведения, «этические нарушения», «правонарушения», «вина» и «ответственность» роботов. Авторы также анализируют антропоморфизм как одну из стратегий социальной робототехники, ее плюсы и минусы.

По мнению авторов, законодательство должно исходить из функций искусственного интеллекта и быть единообразным для всех возникающих его форм. На данном этапе развития для роботов (искусственного интеллекта) следует устанавливать правовой режим, а правовым статусом они смогут обладать только в будущем, когда приобретут реальную автономность и смогут осуществлять собственный свободный выбор поведения.

Ключевые слова: искусственный интеллект, робот, законы робототехники, антропоморфизм, «этика» роботов, «правомерность» поведения роботов, «вина» робота, «ответственность» робота, правовое регулирование, уголовно-правовая оценка.

Цитирование: Батури Ю. М., Полубинская С. В. Искусственный интеллект: правовой статус или правовой режим? // Государство и право. 2022. № 10. С. 141–154.

DOI: 10.31857/S102694520022606-7

ARTIFICIAL INTELLIGENCE: LEGAL STATUS OR LEGAL REGIME?

© 2022 Yu. M. Baturin¹, *, S. V. Polubinskaya², **

¹Lomonosov Moscow State University

²Institute of State and Law of the Russian Academy of Sciences, Moscow

*E-mail: baturin@ihst.ru

**E-mail: svepol@yandex.ru

Received 28.06.2022

Abstract. In Russia, like elsewhere in the world, artificial intelligence, robots and robotic technologies have been introduced into various fields of economy, and state and public life. The variety of spheres of application in which artificial intelligence in different forms is currently used and will be used in future, have implications for the legal system and require adequate legal developments that would take into account the features of these fundamentally novel and previously unknown technologies. The solution of this complex task, in its turn, involves theoretical understanding of multiple conceptual issues that can affect the success or failures of legal regulations in the area of public relations under consideration.

The article considers Asimov's laws of robotics and ethical norms that have been proposed for artificial intelligence in the Concept of development of regulation of relations in the sphere of artificial intelligence and robotics technologies for the period till 2024. The authors discuss the subject of criminal law assessment of robots' behavior, and propose the understanding and definitions of notions like "ethics" and "lawfulness" of robots' behavior, "ethical violations", "offences", "guilt" and "responsibility" of robots. The authors also analyze anthropomorphism as one of the strategies of social robotics, its pros and cons.

The authors conclude that legislation should use as a starting point the functions of artificial intelligence and it should be uniform towards all emerging forms. At present stage of development, it is necessary to have a legal regime established for robots (artificial intelligence), while the legal status for them is the matter of future, when they will acquire real autonomy and will be able to carry out their own free choice of behavior.

Key words: artificial intelligence, robot, Asimov's laws of robotics, anthropomorphism, "ethics" of robot, "lawfulness" of robot's behavior, "guilt" of robot, "responsibility" of robot, legal regulation, criminal law assessment.

For citation: Baturin, Yu. M., Polubinskaya, S.V. (2022). Artificial intelligence: legal status or legal regime? // Gosudarstvo i pravo=State and law, No. 10, pp. 141–154.

Об исходных понятиях. Начнем с рабочих определений. В Национальной стратегии развития искусственного интеллекта на период до 2030 года¹ искусственный интеллект определяется как «комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые как минимум с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений» (п. 1.5а).

Обратим внимание на «поиск решений без заранее заданного алгоритма». Следует добавить: «как правило, без заранее заданного алгоритма», потому что даже человек во многих случаях действует по алгоритмам. (Выходишь на проезжую часть, посмотри налево; дойдешь до середины дороги, посмотри направо — вполне понятный алгоритм, действующий на уровне рефлекса. В странах с левосторонним движением, естественно, наоборот.) С такой поправкой определение искусственного

интеллекта оказывается вполне операциональным, т.е. с ним можно работать².

Теперь — о роботах. Здесь предлагается весьма широкий спектр определений.

Робот — искусственно (т.е. не Природой) созданная машина, проявляющая физические и интеллектуальные способности, принимающая рациональные и иррациональные решения, но не являющаяся живой в биологическом смысле³. Сходное определение дают авторы одной из авторитетных работ по

² О понятии и признаках искусственного интеллекта см.: Модели правового регулирования создания, использования и распространения роботов и систем с искусственным интеллектом / под общ. ред. В.Б. Наумова. СПб., 2019. С. 13–24; Бактеев Д.В. Искусственный интеллект: этико-правовые основы. М., 2021. С. 8–28.

³ В нормативных и программных документах, а также в литературе можно встретить различные определения роботов, включая их виды (см., напр.: «ГОСТ Р 60.0.0.4-2019/ ИСО 8373:2012. Национальный стандарт Российской Федерации. Роботы и робототехнические устройства. Термины и определения», утвержденный и введенный в действие приказом Росстандарта от 14.02.2019 г. № 31-ст // СПС «КонсультантПлюс»; European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL) [Электронный ресурс]. — Режим доступа: URL: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html (дата обращения: 12.05.2022); Bertolini A. Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules // Law, Innovation and Technology. 2013. Vol. 5. Iss. 2. P. 217–219; Jones C.P.A. The Robot Koseki: A Japanese Law Model for Regulating Autonomous Machines // Journal of Business and Technology Law. 2019. Vol. 14. Iss. 2. P. 413–418; Модели правового регулирования создания, использования и распространения роботов и систем с искусственным интеллектом / под общ. ред. В.Б. Наумова. С. 24–32).

¹ См.: Указ Президента РФ от 10.10.2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года») // СПС «КонсультантПлюс».

праву роботов⁴, обращая внимание на то, что «робот — это нечто изготовленное (manufactured), перемещающееся в пространстве, принимающее, похоже, рациональные решения о том, что [ему] следует делать, и [при этом] являющееся машиной»⁵.

Важно иметь в виду разнообразие роботов и иных систем, использующих искусственный интеллект и технологию глубокого обучения (deep learning), в зависимости от сложности их алгоритмов и задач, для решения которых они предлагаются. В этой связи Дж. М. Балкин обращает внимание на то, что «роботы и другие интерактивные объекты не обязательно должны быть спроектированы каким-то определенным образом. А поскольку существует континуум потенциальных конструкций и множество различных потенциальных применений, существует также континуум потенциальных эффектов, связанных с этими новыми технологиями»⁶. Отсюда и возможность использования различных способов и методов их регулирования.

Не будем подробно разбирать достоинства и недостатки приведенных определений роботов. Дадим операциональное (рабочее) определение, которое позднее, вероятно, будет скорректировано: робот — функциональная конструкция, автономный агент, снабженный инструментальными средствами, управляемыми искусственным интеллектом данного агента.

У этого определения есть слабые места (точнее — места, требующие уточнения), хотя они и необходимы для первичной (исходной) дефиниции. Во-первых, относительность понятия «автономный». Так, марсоход (робот) получает команду от человека-оператора: «двигаться к тому валуну», и в этом отношении он не автономен. Но на пути к указанному валуну марсоход вполне автономен: он обязан обнаруживать препятствия и избегать их. Таким образом, «автономность» — переменный параметр: иногда она есть, но в большинстве случаев роботы являются физическим и перцептивным продолжением удаленных операторов (роботы с неочевидной автономностью). Еще более сложный случай, когда наблюдатель не знает, что удаленный оператор — тоже робот. Роботы постепенно становятся все более автономными по мере того, как мы решаем технические проблемы их «восприятия» и «мышления».

⁴ «Робот — это сконструированная система, которая демонстрирует физическое и психическое поведение, но не является живой в биологическом смысле» (см.: *Richards N.M., Smart W.D. How should the law think about robots? // In: Robot Law / R. Calo, A.M. Froomkin, Ia. Kerr (eds). UK, USA, 2016. P. 6).*

⁵ Ibid.

⁶ *Balkin J.M. The Path of Robotics Law // California Law Review Circuit. 2015. Vol. 6. P. 50.*

Во-вторых, следует отметить, что приписывание роботу качества агентности зависит от наблюдателя: робот может казаться обладающим агентностью внешнему наблюдателю, но отнюдь не быть таковым для внутреннего наблюдателя — проектанта (создателя) робота или оператора, дистанционно управляющего им (отсутствие автономности). Боевые дроны или наземные телеуправляемые роботы могут заменить человека в опасных ситуациях: обезвреживание бомбы и т.д. Хотя они полностью контролируются человеком-оператором и не обладают автономными возможностями, они часто выглядят интеллектуальными для внешнего наблюдателя.

Роботы являются и на протяжении многих лет останутся лишь инструментами. Это сложные инструменты, использующие сложное программное обеспечение, но по своей сути они ничем не отличаются от молотка, электродрели или автоматики автомобиля. По мере усиления автономности системы становится все труднее и труднее сформировать связь между входами (командами) и выходами (поведением робота), но она есть и является детерминированной.

Как отмечает Р. Кало, среди исследователей «существует определенный консенсус относительно понимания того, что роботы — это механические объекты, которые воспринимают мир, обрабатывают то, что они воспринимают, и, в свою очередь, действуют в отношении этого мира». Способности «воспринимать — думать — действовать» отличают роботов от других технологий⁷. И потому «роботов лучше всего рассматривать как искусственные объекты или системы, которые воспринимают, обрабатывают и действуют в отношении окружающего мира, по крайней мере в некоторой степени»⁸. Но как роботы воспринимают объекты?

Один и тот же набор входных сигналов будет генерировать один и тот же набор выходов каждый раз. Но ни один датчик не является совершенным, и робот никогда не увидит один и тот же входной сигнал дважды, для него это будет другой сигнал. Никакие две камеры («глаза»), передающие изображения, никогда не дадут тождественные выходные сигналы из-за тонких изменений освещения и ошибок измерения в самой камере. Человек может быть не в состоянии их различать, но программное обеспечение робота видит разницу. Проблема в том, что такое разное поведение, казалось бы, в одинаковых ситуациях может быть истолковано как «свобода воли» или самостоятельность робота.

Для нас наиболее важно, что искусственный интеллект есть часть робота, причем правовая проблема, вынесенная в заголовок статьи, связана

⁷ См.: *Calo R. Robotics and the Lessons of Cyberlaw // California Law Review. 2015. Vol. 103. No. 3. P. 529.*

⁸ Ibid. P. 531.

именно с искусственным интеллектом, а инструментальные средства лишь обеспечивают выполнение заданных функций. Например, космический робот Федор (FEDOR – Final Experimental Demonstration Object Research) не слишком функционален. В самом деле, зачем в невесомости роботу ноги, если ходить нет необходимости. И в то же время лишняя пара рук была весьма полезна⁹. Функциональность робота выражается не только в предназначении для выполнения определенных работ, но в первую очередь определяется заданным функционалом искусственного интеллекта.

О законах робототехники и отношении к ним. Вернемся к искусственному интеллекту как к центральному персонажу нашего обсуждения.

Распоряжением Правительства РФ от 19 августа 2020 г. № 2129-р была утверждена Концепция развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 года¹⁰ (далее – Концепция или Концепция ИИ–2024). В ней приводится важная оговорка: «Термины, используемые в настоящей Концепции, употребляются в значении, приведенном в Национальной стратегии развития искусственного интеллекта на период до 2030 года» (разд. I-1). Правда, этот принцип проводится не вполне последовательно. Так, «алгоритмическая прозрачность систем искусственного интеллекта» (разд. I-4) не слишком соответствует отмеченному признаку определения искусственного интеллекта, который действует «без заранее заданного алгоритма».

Системы искусственного интеллекта, согласно Концепции, характеризуются «неспособностью непосредственно воспринимать этические и правовые нормы, учитывать их при осуществлении каких-либо действий» (разд. I-1). Возникает вопрос: какие нормы имеются в виду – регулирующие поведение человека или поведение искусственного интеллекта, заложенные непосредственно в него? Ответ дает та же Концепция: «Развитие технологий искусственного интеллекта и робототехники должно основываться на базовых этических нормах» (разд. I-3). Следовательно, в основу действий искусственного интеллекта должны быть заложены нормы, по природе своей являющиеся этическими, но это будет этика не людей, а искусственного интеллекта (роботов).

В Концепции ИИ–2024 перечисляются подобные нормы. Некоторые из них очень похожи на знаменитые законы робототехники, сформулированные

Айзеком Азимовым, но в целом это несколько другой комплекс базовых принципов разработки искусственного интеллекта. Сравним их бегло.

Первый закон робототехники А. Азимова. Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред.

Концепция ИИ–2024 налагает запрет на причинение вреда человеку по инициативе систем искусственного интеллекта и робототехники (разд. I-3). Видно, что эта норма существенно отличается от первого закона. В самом деле, робот не может по собственной инициативе причинить вред человеку, но может по приказу другого человека, а также может спокойно наблюдать, как человек подвергается опасности.

Второй закон. Робот должен повиноваться всем приказам, которые дает человек, кроме тех случаев, когда эти приказы противоречат первому закону.

В Концепции ИИ–2024 аналог второго закона сформулирован следующим образом: «Подконтрольность человеку (в той мере, в которой это возможно с учетом требуемой степени автономности систем искусственного интеллекта и робототехники и иных обстоятельств)» (разд. I-3).

Третий закон. Робот должен заботиться о своей безопасности в той мере, в которой это не противоречит первому или второму законам (в Концепции ИИ–2024 аналога нет).

Позже А. Азимов добавил *четвертый закон* (известный как закон Зеро): ни один робот не может причинить вред человечеству или посредством бездействия позволить человечеству пострадать (в Концепции ИИ–2024 также аналога нет)¹¹.

Не только логические несоответствия обнаруживаются в нормах Концепции ИИ–2024 при самом беглом анализе, но оказывается, что и ставший классическим набор законов Азимова тоже внутренне противоречив. Дело в том, что все эти законы-нормы сформулированы без обращения к базовому уровню – метавыбору, связанному с отношением к добру и злу, т.е. к этическим основам¹². Поэтому специалистами обсуждается вопрос о необходимости заложить в роботов при

¹¹ См.: Lichocki P., Billard A., Kahn P.H. The Ethical Landscape of Robotics // IEEE Robotics&Automation Magazine. 2011. Vol. 18. Iss. 1. P. 46.

¹² См.: Батури́н Ю.М. Гамлет как робот: основы правового регулирования искусственного интеллекта надо искать в хорошо забытом прошлом советской юридической науки: материалы Четвертой Междунар. науч.-практ. конф. «Бачиловские чтения» (5–6 февраля 2021 г.) / отв. ред. Т.А. Полякова, А.В. Минбалеев, В.Б. Наумов. Саратов, 2022. С. 265–274; см. также: Полякова Т.А., Минбалеев А.В., Кроткова Н.В. Развитие науки информационного права и правового обеспечения информационно-безопасности: формирование научной школы информационного права (прошлое и будущее) // Государство и право. 2021. № 12. С. 102, 103. DOI: 10.31857/S102694520017761-8

⁹ См.: Фото робота Федора в полный рост [Электронный ресурс]. – Режим доступа: URL: https://yandex.ru/images/search?pos=34&img_url=https%3A%2F%2Fa.dcd.net%2F2cc-5751s1920.jpg&text=космический%20робот%20федор%20фото-&lr=213&grpt=simage&source=segr (дата обращения: 15.06.2022).

¹⁰ См.: СПС «КонсультантПлюс».

их создании определенные этические принципы. И хотя большинство специалистов говорят о необходимости этического искусственного интеллекта, согласия, каким же именно принципам должно следовать поведение его носителей, пока нет.

Итак, наше внимание будет направлено в первую очередь на «этику» роботов, основное назначение которой предотвратить их поведение, причиняющее вред людям, имуществу, окружающей среде, другим роботам и самим себе, наконец. В литературе отмечается, что законы робототехники «основаны на практической морали, которая исходит из предположения, что роботы обладают достаточными качествами для действия и когнитивными способностями для принятия моральных решений»¹³. Отсюда вывод об отсутствии их практической применимости для регулирования поведения роботов. Поэтому в резолюции Европейского парламента «Нормы гражданского права о робототехнике» прямо указано, что «законы Азимова должны рассматриваться как направленные на разработчиков, производителей и операторов роботов, включая роботов, наделенных встроенной автономией и самообучением, поскольку эти законы не могут быть преобразованы в машинный код»¹⁴. С этими утверждениями можно поспорить. Для того чтобы признать робота субъектом, чье поведение управляется нормами морали, и, соответственно, вести речь о его возможной ответственности за причиненный вред, необходимо, чтобы он обладал автономностью, т.е. способностью «выполнять задачи по назначению на основе текущего состояния и восприятия внешней среды без вмешательства человека»¹⁵, и мог сформировать намерение причинить вред¹⁶. Заметим, что эти условия выведены фактически из оснований юридической ответственности для человека. Человек, обладающий сознанием и волей, осуществляет выбор вариантов поведения и подлежит, к примеру, уголовной ответственности, когда действует виновно по отношению к совершенному общественно опасному деянию и его общественно опасным последствиям (ст. 5 УК РФ).

Об уголовно-правовой оценке поведения роботов. В литературе описывается по преимуществу три

возможные ситуации, в которых поведение роботов может иметь уголовно-правовое значение: 1) робот является орудием преступления и используется с этой целью другим человеком (сюда можно отнести случаи, когда робот специально программируется для совершения преступления); 2) робот совершает «преступление» из-за ошибок в программировании или недостатков, допущенных при изготовлении; 3) робот «совершает» преступление, так сказать, по собственной воле¹⁷. В первых двух случаях субъектом преступления является человек, действующий в первом случае умышленно, а во втором по неосторожности (небрежности, как правило).

Для третьего случая необходимо, чтобы робот был признан субъектом права, сходным с человеком¹⁸, и обладал необходимыми когнитивными способностями, чтобы нести уголовную ответственность, а в его деянии должны быть установлены все объективные и субъективные признаки соответствующего состава преступления¹⁹. Г. Халлеви называет эту ситуацию «моделью прямой [уголовной] ответственности» искусственного интеллекта. По его мнению, она не отличается от уголовной ответственности человека, и имеющиеся в распоряжении уголовного права наказания вполне применимы и к искусственному интеллекту. Надо лишь создать нового субъекта для уголовного права. В своих рассуждениях он использует аналогию искусственного интеллекта с юридическими лицами, уголовная ответственность которых существует в целом ряде зарубежных государств²⁰.

Другие авторы отмечают, что на данном этапе развития мы не можем признать носителей искусственного интеллекта равным людям, как и наличие у таких объектов качеств, необходимых для возложения на них уголовной ответственности, однако не исключают этого в будущем²¹.

В литературе также предлагается разработать уголовный кодекс для умных роботов, управляемых алгоритмами, позволяющими принимать «нетривиальные морально значимые решения»,

¹³ Murphy R., Woods D.D. Beyond Asimov: The Three Laws of Responsible Robotics // IEEE Intelligent Systems. 2009. Vol. 24. No. 4. P. 14.

¹⁴ European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

¹⁵ См.: ГОСТ Р 60.0.0.4-2019/ИСО 8373:2012. Национальный стандарт Российской Федерации. Роботы и робототехнические устройства. Термины и определения (п. 2.2).

¹⁶ См. подр.: Sullins J.P. When Is a Robot a Moral Agent? // International Review of Information Ethics. 2006. Vol. 6. P. 28, 29.

¹⁷ См.: Hallevy G. The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control // Akron Intellectual Property Journal. 2010. Vol. 4. Iss. 2. P. 174; Simmler M., Markwalder N. Guilty Robots? – Rethinking the Nature of Culpability and Legal Personhood in the Age of Artificial Intelligence // Criminal Law Forum. 2019. Vol. 30. Iss. 1. P. 8.

¹⁸ О правосубъектности искусственного интеллекта см.: Модели правового регулирования создания, использования и распространения роботов и систем с искусственным интеллектом / под общ. ред. В.Б. Наумова. С. 53–67; Бахмев Д.В. Указ. соч. С. 118–124.

¹⁹ См.: Hallevy G. Op. cit. P. 186, 187.

²⁰ См. подр.: ibid. P. 186–199.

²¹ См., напр.: Simmler M., Markwalder N. Op. cit. P. 20, 27.

«способных сообщать о своих решениях людям» и имеющих возможность действовать «без непосредственного контроля со стороны человека»²². Такие характеристики робота, с точки зрения Ю. Ху, являются необходимыми предварительными условиями для начала разговора об уголовной ответственности роботов. При этом он отмечает, что «если нет веских причин для введения уголовной ответственности для роботов, мы должны воздерживаться от этого»²³.

В большинстве случаев робот рассматривается как гуманоид, как вид человека. Мы видим, как срабатывает первая и особенно соблазнительная метафора для роботов, которую следовало бы отвергнуть с порога — идея о том, что роботы «такие же, как люди». Ловушка антропоморфного (гуманоидного) робота приведет нас к ложным предположениям о возможностях роботов и подталкивает воспринимать их как нечто большее, чем просто роботы, даже если мы изо всех сил будем стараться этого не делать, а это, в свою очередь, повлечет появление плохого законодательства. Антропоморфная проекция человеческих качеств опасна

новую технологию как новую форму одной из предшествующих. Если бы аналогия использовалась строго логически, так чтобы в ее структуре переносимый предикат представлял бы собой «фиксированное свойство “быть истинным”»²⁶, то все было бы хорошо. Однако на практике аналогия становится простым эвристическим приемом, в котором превращается в метафору. И здесь многое меняется, в том числе в мышлении юриста, поскольку включаются другие когнитивные структуры личности (установлено, что за метафоры отвечает совсем другой участок мозга, чем, скажем, за метонимию), создающей некую концепцию, представленную в виде текста²⁷.

Внешне изменения минимальны и совершенно незаметны. По определению, идущему еще от Аристотеля, метафора (др.-греч. *μεταφορά* — «перенос», «переносное значение») — свернутое (не названное) сравнение одного объекта с другим по некоторому общему признаку. Метафора получается из сравнения элиминированием сравнительного слова «как».

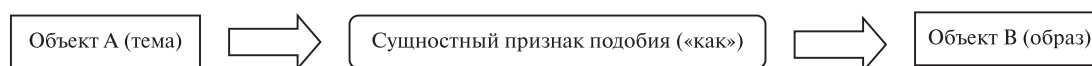


Рис. 1. Простейшая структура метафоры

при попытках разработать законодательство для роботов. Кроме того, перенесение человеческих качеств и способностей на искусственный интеллект способно породить ожидания «поведенческой и когнитивной сложности, которая может оказаться нереализуемой»²⁴. Особенно опасен перенос подобных ожиданий на нормативное регулирование искусственного интеллекта.

О юридических метафорах. Юристов с первого дня обучения в университете учат юридическому мышлению, в котором важное место занимает аналогия (права и закона) как средство «отображения и развития наших знаний путем индуктивного обобщения связей между рядом признаков, сходных у двух или более предметов, отношений и признаков, присущих одному из них»²⁵. В череде быстро сменяющихся друг друга информационных технологий закон почти всегда рассматривает

Сопоставляемый объект А по аналогии представляется в образе В, а признак подобия «как» убирается.

Ничего необычного для использования метафоры в науке нет. Математика, физика, другие естественные и технические науки полны метафор, начиная от «физической картины мира», включая «очарованные частицы», «черный ящик» и заканчивая «демоном Максвелла» и «котом Шредингера». Понятно, что мир никто не рисовал на картине, а кот Шредингера не имеет отношения к семейству кошачьих. Задача метафор — через смысловой образ подтолкнуть мысль к объяснению явления.

Но при осмыслении новых технологий в юридических понятиях метафоры, которые мы используем для их понимания, могут стать опасными. Если мы неправильно будем использовать метафоры для искусственного интеллекта и роботов, то можем положить начало разветвленной сети причинно-следственных связей, приводящей к ошибкам,

²² См.: Hu Y. Robot Criminals // University of Michigan Journal of Law Reform. 2019. Vol. 52. Iss. 2. P. 490.

²³ Ibid. P. 492.

²⁴ Duffy B. R. Anthropomorphism and the social robot // Robotics and Autonomous Systems. 2003. Vol. 42. P. 183.

²⁵ Петровский Н. А. Аналогии в юриспруденции и правоприменительной практике // Вестник Адыгейского гос. ун-та. 2005. Вып. 3. С. 128.

²⁶ Там же. С. 131.

²⁷ См.: Шевченко Л. Л. Метафора как средство моделирования концептуальной системы автора: дис. ... канд. филол. наук [Электронный ресурс]. — Режим доступа: URL: <http://www.dslib.net/germanskie-jazyki/metafora-kak-sredstvo-modelirovaniia-konceptualnoj-sistemy-avtora.html> (дата обращения: 15.06.2022).

исправить которые будет уже очень трудно²⁸. Опыт информационного права, компьютерного права и других правовых отраслей, подверженных влиянию технологий, преподает важный урок для технологически сложного применения права — особенно важно использовать правильную аналогию (метафору) для новой технологии, когда дело касается принципиально новых и неизвестных ранее технологий. То, как мы будем регулировать поведение роботов (и искусственного интеллекта), будет зависеть от аналогий (метафор), которые мы используем, думая о них. Существует множество конкурирующих метафор для различных видов роботов, и правильное понимание метафоры будет иметь чрезвычайно важные последствия для успеха или неудачи так или иначе введенного закона (или законов) робототехники.

Инженерное дело и юриспруденция имеют много общего. И то и другое требует тщательной оценки пределов применения своего продукта, сравнивает затраты и выгоды, устанавливает причинно-следственные связи. И то и другое использует схожие концепции и термины, например, «управление» и «регулирование». И свои метафоры. Для права важно, хотя и трудно, отличать имеющее ясный смысл и определенное содержание техническое понятие от технической метафоры. А в области искусственного интеллекта и робототехники инженерных метафор более чем достаточно.

Чтобы перейти от метафор к содержательным понятиям, мы должны начать с того, чтобы построить или спроектировать упрощенный образец (модель) искусственного интеллекта, обладающего внутренним миром и способностью генерировать субъективные намерения. Таким образом, и свобода воли, и свобода искусственного интеллекта (робота) совершать намеренные поступки должны включаться в модель в виде отдельных переменных (параметров). Интеллект — это рациональность, включая и моральный выбор, который осуществляется особым процессором, «совестью», регулирующим поступки искусственного интеллекта и заставляющим его в случае отклонения от «правильного» выбора подвергнуть себя «моральным» терзаниям, т.е. испытывать чувство «вины». Это значит, что мы можем говорить об «этике» искусственного интеллекта.

Об «этике», «вине» и «ответственности» роботов. Что же такое «этика» робота? (Здесь и далее в данном разделе кавычки означают, что мы используем метафору до тех пор, пока не выясним ее смысл и не введем соответствующее понятие.) Ясно, что это набор внутренних (заложенных проектиантами) правил (ограничений), за нарушение которых

робот с искусственным интеллектом ответствен перед собой. Что означает «перед собой»? Смысл простой: нарушение внутренних (заложенных) правил влечет внутреннюю же реакцию робота на нарушение. Следовательно, этическое поведение (действия) робота — режим обратной связи в его управляющем звене — искусственном интеллекте. Таким образом, этика робота (искусственного интеллекта) — это система заложенных в контур управления обратных связей (положительных и отрицательных).

Что такое «правомерность» поведения (действий) робота? Это выбор поведения (действий) робота, управляемого искусственным интеллектом, не нарушающего внешние, заложенные в него проектантом (создателем) правила — юридические и технические, выраженные в алгоритмах нормы, и, следовательно, не требующие по этой причине приостановки его функционирования. Соответственно, «правомерное» поведение (действия) робота — режим обратной связи в его управляющем звене — искусственном интеллекте, не ведущий к приостановке его действий.

Что такое нормативные ограничения поведения (действий) робота? Внутренние, заложенные в искусственный интеллект правила суть нормативные ограничения действий робота, нарушение которых влечет его самостоятельную реакцию. Но существуют и нормативные ограничения (правила), за нарушение которых искусственный интеллект «ответствен» перед внешним миром (миром людей). И то и другое — нормативные ограничения поведения робота. Выход за эти нормативные ограничения означают, соответственно, «неэтичность» и неправомерность поведения робота.

Нарушение внешних нормативных ограничений — это «вина» робота. Точнее — «осознание» искусственным интеллектом своего нарушения. «Вина» робота логически формулируется на основании введенной аксиоматики²⁹, т.е. формы «вины» робота могут быть разными, но при введении аксиоматики желательно не попасть в ловушку антропоморфизма, который тут же уведет нас в сторону от сути.

Иллюстративный пример, который не следует смешивать с предлагаемой для «вины» роботов работающей аксиоматикой ввиду того, что ей должны предшествовать достаточно развитый логический анализ и сравнение вариантов.

Обозначим через N одноместный лингвистический оператор «вина», а через $\tilde{N}$ — двухместный оператор «быть виновным». Строчными латинскими буквами $\{a, b, c, \dots\}$ обозначим некоторые действия,

²⁸ См.: Richards N.M., Smart W.D. Op. cit. P. 16–18.

²⁹ См.: Влияние научно-технического прогресса на юридическую жизнь. М., 1988. С. 114.

греческими буквами $\{\alpha, \beta, \gamma, \dots\}$ — определенные условия. Тогда $N(b)$ означает «действие b виновное», а $\tilde{N}(\gamma, b)$ — «действие b при условии γ виновное». Условия $\{\alpha, \beta, \gamma, \dots\}$ могут быть самого различного характера. Например, робот предпринимает действия, направленные на предотвращение гибели человека, вычислив, что эти действия являются достаточными, но ему не удастся спасти человека.

Примем обычные логические обозначения (кванторы): $\neg$ — отрицание, $\wedge$ — конъюнкция («и»), $\vee$ — дизъюнкция («или»), $\rightarrow$ — импликация («если... то...»), $\exists$ — квантор существования. Попытаемся ввести исходную аксиоматику для описания структуры «вины» робота (или по крайней мере части этой структуры).

Аксиома 1. $\neg(b) \rightarrow (N(b) \vee N(\neg(b)))$

(Своего рода отрицание правила исключенного третьего применительно к «вине» робота: ни какое-либо действие, ни соответствующее бездействие не являются с необходимостью виновными).

Аксиома 2. $N(a) \equiv \exists \alpha (\tilde{N}(\alpha, a) \wedge \alpha)$

(Необходимое и достаточное условие «виновного» действия состоит в том, что существует условие «виновности», и это условие выполняется. Иными словами, « α есть условие «виновности» a , и α — истинно»).

Аксиома 3. $N(b) \rightarrow N(\beta \wedge b)$

(Если совершение действия b является виновным, то оно обусловлено каким-то β). И так далее, до полного аксиоматического описания структуры «вины» робота.

Из структуры «вины» робота логически выводится понятие «невиновности» робота и др. На основе логики «невиновности» создается логическое устройство (логический компаратор), сравнивающий логику «виновного» действия искусственного интеллекта с логикой «невиновности». Искусственный интеллект заносит «виновные» действия в каталог нарушений, база которого закладывается проектантами робота.

Нарушения делятся на «этические» и «правонарушения» робота. «Этическим» называется виновное нарушение, не приведшее к существенному ущербу здоровью и / или собственности людей и / или к их гибели, которое искусственный интеллект способен скорректировать самостоятельно (самообучение). «Ответственность» за «этическое нарушение» состоит в том, что искусственный интеллект использует факт нарушения и анализ этого факта для самообучения, которое сводится к исключению действий, занесенных в каталог «этических нарушений», из последующего поведения робота с целью постепенного снижения частоты «этических нарушений».

Что такое «юридическая ответственность» робота (искусственного интеллекта)? Внешняя (осуществляемая извне искусственного интеллекта) приостановка действий робота с возможной их коррекцией или прекращением его эксплуатации.

«Правонарушением» робота будем называть «виновное» нарушение, приведшее к существенному ущербу здоровью и / или собственности людей и / или к их гибели, которое искусственный интеллект не способен скорректировать самостоятельно в процессе самообучения, и требующее вмешательства человека следующих типов:

вмешательство «медицинского» характера, своего рода «лечение» искусственного интеллекта, коррекция поведения робота, которую должны осуществлять проектанты, разработчики, изготовители робота (искусственного интеллекта);

вмешательство правоохранительных органов, выносящих решение о внешней приостановке действий искусственного интеллекта, о прекращении эксплуатации робота данной модели до завершения расследования «правонарушения» или отзыва модели для модификации или полного снятия модели с производства.

Мы использовали термины «этика», «юридическая ответственность», «вина», «виновность», «невиновность», «правонарушение» и др. в кавычках, поскольку речь шла всего лишь о некоторых аналогах понятий, применимых к миру людей. По сути, «вина» искусственного интеллекта и по конструкции, и по способам ее установления — это совсем не вина человека. «Правонарушение», совершенное роботом, не есть нарушение норм объективного права, действующего в мире людей. Поэтому желательно во избежание путаницы ввести для обсуждения темы нормативного регулирования отношений по поводу создания и использования искусственного интеллекта специальную терминологию:

«этика» искусственного интеллекта — установленные (имеющие управленческий генезис) обыкновения (usage) искусственного интеллекта, регуляторный образ (модель, гомоморфное отображение) закладываемого в искусственный интеллект поведения робота, или просто — *обыкновения* искусственного интеллекта;

«этическое нарушение» — внутренний нормативный сбой искусственного интеллекта;

«вина» искусственного интеллекта — аномалия его функционирования;

«виновность» искусственного интеллекта — «осознание» (внутренний контроль) аномальности поведения (действий) робота;

«невиновность» искусственного интеллекта — нормальность поведения (действий), норма, адекватность робота;

«виновные действия» — аномальные действия (поведение) робота;

«правонарушение» искусственного интеллекта — дефект действия, отклонение действия от установленной нормы поведения робота (выход за пределы допуска вариабельности какого-либо параметра);

«юридическая ответственность» искусственного интеллекта — нормативный отклик государства на выход параметров функционирования робота за пределы допуска (действия, которые предписано выполнить в отношении робота);

вмешательство «медицинского» характера — коррекция поведения искусственного интеллекта согласно заданным техническим нормам и техническим условиям;

вмешательство правоохранительных органов — возложение юридической ответственности на проектантов, производителей, собственников, пользователей роботов в соответствии с обстоятельствами возникновения дефекта поведения робота (только в данном случае понятие «юридическая ответственность» используется в прямом смысле слова применительно к будущему законодательству о создании и использовании систем искусственного интеллекта).

Закладываемые в искусственный интеллект и роботов нормы поведения в машинном коде требуют предельной точности и строгости, большей, чем в действующем законодательстве. Это заставляет проектантов, производителей, собственников и пользователей роботизированных систем идти на повышенные риски.

Форма тоже имеет значение. Итак, роботы еще не достигли того уровня возможностей, который ассоциируется у публики с научной фантастикой, но они двигаются к этому. «Роботы» в обыденном понимании (программируемые автоматы) существуют уже сегодня. Роботы используются в торговых центрах, медицинских учреждениях, аэропортах. Роботы убирают пылесосом дом. Компьютерные алгоритмы, которые управляют ими, просты, но, похоже, с их помощью принимаются рациональные решения. Роботы передвигаются по полу, избегая посторонних объектов в загроможденной реальной среде. Это уже впечатляющее достижение. Роботам больше не нужен сопровождающий, выполняющий функции надсмотрщика, переводчика, механика и телохранителя.

Роботы начинают взаимодействовать с людьми, которые ничего не знают о роботах. Люди, не имевшие ранее опыта общения с роботами, теперь сталкиваются с ними и сотрудничают с ними

напрямую. Хотя это необходимо, это также усложняет ситуацию в огромной степени. Люди непредсказуемы, им легко повредить и трудно угодить³⁰.

То, как мы думаем, понимаем и концептуализируем роботов и искусственный интеллект, будет иметь реальные последствия на стадии разработки концепции, проектирования, законодательства и потребительского спроса. Люди по-разному реагируют на технологии, которые имеют (или не имеют) антропоморфную форму, глаза, как у человека, говорят человеческим голосом (а не металлическим машинным «голосом»). Эти поведенческие реакции, по-видимому, жестко запрограммированы, наглядное преимущество человекоподобных по сравнению с нечеловекоподобными технологиями будет оказывать реальное влияние на то, как потребители будут реагировать на роботов в своих домах, их принятие и доверие к ним. Поэтому мы делаем из «ловушки антропоморфизма» исключение, связанное именно с желаниями человека-потребителя.

Не только законодатели подвержены антропоморфным воздействиям, но и общество, для которого пишутся законы. Оно тоже будет склонно антропоморфизировать роботов с (предполагаемыми) человеческими характеристиками. Хотим ли мы, чтобы роботы походили на домработниц, соседей или детей? Или заменяли нам домашних животных? От ответов на эти вопросы зависит не только, какие функции будут в роботов заложены, но и то, как они физически выполнены. Домработницы, соседи и дети обычно на нас не нападают; а некоторые домашние животные могут это сделать. Дети и домашние животные менее автономны, чем домработницы и соседи, а дети, соседи и домработницы (но не домашние животные) антропоморфны. Список прообразов робота легко может быть продолжен³¹.

Антропоморфизм представляет собой широко распространенный феномен перенесения человеческих качеств (разума, эмоций, намерений, мотивов и т.д.) на живые и неживые объекты окружающего мира. Люди переносят свою способность к сознательному поведению и эмоциональному взаимодействию на домашних питомцев, свои автомобили и компьютеры, нередко общаясь с ними как с себе подобными³².

В социальной робототехнике для производства роботов, способных «вступать в значимое социальное

³⁰ См.: Richards N.M., Smart W.D. Op. cit. P. 9, 10.

³¹ См.: ibid. P. 16, 17.

³² О психологических процессах, лежащих в основе антропоморфизма, см.: Epley N., Waytz A., Cacioppo J.T. On Seeing Human: A Three-Factor Theory of Anthropomorphism // Psychological Review. 2007. Vol. 114. No. 4. P. 864–886.

взаимодействии с людьми»³³, антропоморфизм имеет практическое значение. Он является стратегией «интеграции роботов в социальное окружение человека»³⁴ и, по мнению Б.Р. Даффи, был неизбежен³⁵.

Антропоморфизм социальных роботов основывается на теоретическом представлении о том, что «чем более подобен человеку носитель [искусственного интеллекта], тем больше вероятность того, что люди будут доверять [ему] и принимать его»³⁶. Результатом такого представления служит наделение подобных объектов способностями передвигаться и говорить. Как замечает Дж. М. Балкин, «люди ассоциируют способность к речи с человечностью в большей степени, чем способность к движению»³⁷.

Эмпирические исследования показывают, что наличие у социальных роботов более близких к человеческим внешним признакам положительно влияет на их восприятие, повышает доверие и облегчает взаимодействие пользователей с ними³⁸. Однако не вполне ясно, до какой степени возможна антропоморфизация робота. Еще в 1970 г. японский робототехник М. Мори выдвинул гипотезу о связи подобия человеку с восприятием привлекательности робота³⁹. Он показал, что сходство с человеком привлекательно до определенного предела, после которого искусственный объект «начинает вызывать у людей резкое когнитивное и эмоциональное отторжение»⁴⁰. М. Мори назвал этот эффект «зловещей долиной» (uncanny valley)⁴¹.

³³ Duffy B.R. Op. cit. P. 178; см. также: Зильберман Н.Н., Стефанцова М.А. Социальный робот: подходы к определению понятия // Современные исследования социальных проблем (электрон. науч. ж-л). 2016. № 11 (67). С. 297–312.

³⁴ Coeckelbergh M. Three Responses to Anthropomorphism in Social Robotics: Towards a Critical, Relational, and Hermeneutic Approach // International Journal of Social Robotics. 2021. P. 1 [Электронный ресурс]. — Режим доступа: URL: <https://doi.org/10.1007/s12369-021-00770-0> (дата обращения: 27.05.2022).

³⁵ См.: Duffy B.R. Op. cit. P. 181.

³⁶ Lockey S., Gillespie N., Holm D. et al. A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions // Proceedings of the 54th Hawaii International Conference on System Sciences. 2021. P. 5466.

³⁷ Balkin J.M. Op. cit. P. 56; см. также: Середкина Е.В. Философские основания прикладного антропоморфизма в социальной робототехнике // Технологос. 2020. № 4. С. 56–63.

³⁸ См., напр.: Belanche D., Casaló L.V., Flavián C. et al. Service robot implementation: a theoretical framework and research agenda // The Service Industries Journal. 2020. Vol. 40. Iss. 3–4. P. 208, 209.

³⁹ См.: Mori M. The Uncanny Valley [Электронный ресурс]. — Режим доступа: URL: <https://erikbuys.files.wordpress.com/2014/02/the-uncanny-valley.pdf> (дата обращения: 06.06.2022).

⁴⁰ Столбова Н.В., Середкина Е.В., Мышкин О.С. Насколько «Зловещая долина» зловеща на самом деле? Опыт деконструкции дискурса // Вестник Пермского ун-та. Философия. Психология. Социология. 2022. Вып. 1. С. 94.

⁴¹ См. подр.: Катерный И.В. Каузальные объяснения эффекта «зловещей долины» в робототехнике: теории и исследовательские данные // Качество и жизнь. 2017. № 4 (16). С. 88, 89.

С развитием робототехники возник интерес к изучению феномена «зловещей долины», но полученные к настоящему времени данные весьма разноречивы. Одни исследователи подтверждают этот эффект, другие, напротив, его не обнаруживают⁴². В этой связи предлагается проверять существование этого феномена в условиях реальных взаимодействий роботов и людей⁴³.

Антропоморфизм относится не только к физической, внешней форме носителей искусственного интеллекта, но и к проходящими в них внутренним процессам. Наделение роботов индивидуальностью, разумом и присущими человеку способностями распространено в произведениях литературы и киноискусства, влияющих на общественные представления о роботах. Достаточно вспомнить мальчика-робота из повестей Е. Велтистова и снятого по их мотивам детского кинофильма «Приключения Электроника», а также лейтенанта-коммандера Дейту из американской фантастической саги «Звездный путь» (Star Trek).

В литературе отмечается ряд проблем, связанных с очеловечиванием искусственного интеллекта⁴⁴. Так, взаимодействие с такими объектами может вызывать у пользователей эмоциональную привязанность, способную влиять на контакты с другими людьми. В июле 2016 г. в японском буддистском храме была проведена заупокойная служба по сломавшимся робособакам AIBO, выпускавшимся компанией Sony с 1999 по 2006 г. В 2014 г. производство запасных частей к AIBO было прекращено, ремонтные центры закрыты. Их владельцы не имели возможности починить механических питомцев, им оставалось только с ними попрощаться по-человечески⁴⁵.

⁴² См., напр.: MacDorman K.F., Green R.D., Ho Ch.-Ch. et al. Too real for comfort? Uncanny responses to computer generated faces // Computers in Human Behavior. 2009. Vol. 25. Iss. 3. P. 695–710; Tinwell A., Grimshaw M., Abdel Nabi D., et al. Facial expression of emotion and perception of the Uncanny Valley in virtual characters // Computers in Human Behavior. 2011. Vol. 27. Iss. 2. P. 741–749; Złotowski J., Proudfoot D., Yogeewaran K. et al. Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction // International Journal of Social Robotics. 2015. Vol. 7. Iss. 3. P. 353; Laakasuo M., Palomäki J., Köbis N. Moral Uncanny Valley: A Robot's Appearance Moderates How its Decisions are Judged // International Journal of Social Robotics. 2021. Vol. 13. Iss. 7. P. 1679–1688; Филимонов В.А., Чернявская В.С. Катастрофы субъективных оценок и их модели: к вопросу инструментализации диагностики эффекта «зловещей долины» // Территория новых возможностей. Вестник Владивостокского гос. ун-та экономики и сервиса. 2021. Т. 13. № 2. С. 170–178.

⁴³ См.: Złotowski J., Proudfoot D., Yogeewaran K. et al. Op. cit. P. 353.

⁴⁴ См., напр.: Salles A., Evers K., Farisco M. Anthropomorphism in AI // AJOB Neuroscience. 2020. Vol. 11. Iss. 2. P. 90, 91; Kaminski M.E., Rueben M., Smart W.D., et al. Averting Robot Eyes // Maryland Law Review. 2017. Vol. 76. Iss. 4. P. 983–1025.

⁴⁵ См.: Remembering AIBO. Nippon. Science. Technology. Feb. 2, 2017 [Электронный ресурс]. — Режим доступа: URL: <https://www.nippon.com/en/views/b00909/> (дата обращения: 02.06.2022).

Кроме того, присутствие в доме искусственного друга создает опасность разглашения информации персонального характера, не только которую он воспринимает и регистрирует, но и сообщенной пользователем в «дружеских беседах». Без ведома владельца домашнего робота информация может попасть к удаленным операторам робота или третьим лицам, не говоря уже о возможностях кибератак.

Как выстраивать законодательство. Но все же мы должны формировать законодательство, основываясь на функциях робота, а не на его форме, что стало бы серьезной ошибкой. Когда мы встраиваем искусственный интеллект под капот автомобиля, он становится частью автомобиля, и соответствующие отношения законодательно регулируются как возникающие по поводу автомобиля. Мы законодательно закрепляем форму, а не функцию. Если же мы автомобиль (салон, колеса и т.д.) встраиваем в робота с искусственным интеллектом, все происходит наоборот. Одни и те же датчики, одно и то же программное обеспечение приводят к одному и тому же результату, но мы разделили его на два разных с юридической точки зрения случая⁴⁶. И эта проблема не единственная среди возникающих в связи с роботизацией. Поскольку роботы становятся все более автономными, вопрос о том, на кого возлагается юридическая ответственность, когда что-то пошло не так, становится сложным. Несет ли ее проектант, производитель, собственник, пользователь (который дал неточную команду) или какая-то другая комбинация субъектов? Ответ на этот вопрос будет еще сложнее в системах, которые автономны в одно время и телеуправляемы в другое, так как в этом случае появляется удаленный оператор, который, возможно, управляет роботом из лучших побуждений, но с ограниченной информацией. Поскольку роботы становятся все более и более многоцелевыми (полифункциональными), будет все труднее априори представить, как они будут использоваться, и, следовательно, сложнее создать всеобъемлющую правовую защиту для них и от них.

Огромное разнообразие сфер применения, в которых могут и будут использоваться искусственный интеллект и роботы, окажет воздействие на правовую систему в самых разных отраслях, включая конституционное, уголовное и гражданское право, защиту прав потребителей, неприкосновенность частной жизни и др.

Мы вступаем в эру, когда искусственный интеллект и роботы будут запрограммированы на выполнение или соблюдение различных норм. Как юридически все это формулировать? Нет в окружаю-

щем мире других устройств, похожих на роботов и искусственный интеллект, поэтому мы должны придумывать подходящие метафоры, но затем переводить их в новые содержательные понятия.

* * *

В юридической традиции и в системологии субъект и объект (предмет) противопоставлены друг другу. В отличие от объекта субъект может быть и причиной, и источником своей собственной активности. Простые автоматы (программы) рассматриваются как объекты. Но с переходом к искусственному интеллекту (Концепция ИИ–2024) объект становится источником своей активности, т.е. приобретает черты субъекта. Назовем такого субъекта «субъект-робот». Чем отличаются «субъект-человек» от «субъекта-робота»? Человек может осуждать других людей и себя за некоторые действия. Робот как объект не может отличать свои добрые или несущие зло поступки, например, от естественных благоприятных погодных условий и природных катаклизмов, т.е. для робота любая причина, в том числе и им порожденная, лежит вне его. Когда мы говорим о «субъекте-человеке» или о «субъекте-роботе», предполагаем, что они обладают свободой выбора. Это означает, что субъект способен совершить именно то, что «хочет» совершить, т.е. интенция «хочет» трансформируется в соответствующую готовность. Наши субъекты обладают свободой выбора при условии, что среда склоняет их к совершению добра или зла, и субъекту известно это. В достаточно грубом приближении «свобода выбора» означает, что при некоторых условиях субъект способен совершить именно то, что он хочет совершить. Человек, а в рамках развиваемой концепции и искусственный интеллект способны строить программы своего будущего поведения. Таким образом, для роботов (искусственного интеллекта) мы устанавливаем правовой режим, а правовым статусом они смогут обладать только в будущем, когда при реальной автономности они смогут осуществлять собственный свободный выбор поведения. Это будущее достаточно далеко, но мы к нему неуклонно движемся. Как заметил Дж. П. Саллинс, «ни один робот в реальном мире или в ближайшем будущем не обладает или не будет обладать такими же когнитивными способностями, как собака-поводырь. Но даже при тех скромных возможностях, которыми обладают роботы сегодня, некоторые из них имеют больше общего с собакой-поводырем, чем с молотком»⁴⁷. То есть уже сейчас в них есть не только инструментальная часть, но и интеллектуальная составляющая, которую мы, сильно забегая вперед, называем искусственным интеллектом.

⁴⁶ См.: Richards N.M., Smart W.D. Op. cit. P. 19.

⁴⁷ Sullins J.P. Op. cit. P. 25.

Выводы:

1. Необходимо профессиональное обсуждение определений робота, искусственного интеллекта и других сходных понятий.

2. Искусственный интеллект, роботы и роботизированные технологии будут быстро проникать в нашу жизнь в самых разных формах. Но наше законодательство должно быть единообразным для всех возникающих форм и исходить из функций системы.

3. Нельзя создавать систему правового регулирования деятельности объектов, принципы действия которых мы не понимаем. Прежде чем создавать законодательство в надежде, что оно будет эффективным, мы должны понять, на что способны роботы (искусственный интеллект), что они еще не могут делать, и что они *никогда* не смогут делать. «Никогда» — это, конечно, тоже метафора длительного периода после того, как мы покинем этот роботизированный мир со своим пониманием искусственного интеллекта.

4. Мы должны использовать наш значительный опыт в области информационного права, посмотреть, как оно проводило аналогии с существующими технологиями и действующим законодательством, где оно преуспело, а где потерпело неудачу. Это поможет обосновать наш выбор метафор и новой терминологии для искусственного интеллекта, роботов и роботизированных технологий наряду с выбором инструментов регулирования.

СПИСОК ЛИТЕРАТУРЫ

1. Батури́н Ю. М. Гамлет как робот: основы правового регулирования искусственного интеллекта надо искать в хорошо забытом прошлом советской юридической науки: материалы Четвертой Междунар. науч.-практ. конф. «Бачиловские чтения» (5–6 февраля 2021 г.) / отв. ред. Т. А. Полякова, А. В. Минбале́ев, В. Б. Наумов. Саратов, 2022. С. 265–274.
2. Бахтеев Д. В. Искусственный интеллект: этико-правовые основы. М., 2021. С. 8–28, 118–124.
3. Влияние научно-технического прогресса на юридическую жизнь. М., 1988. С. 114.
4. Зильберман Н. Н., Стефанцова М. А. Социальный робот: подходы к определению понятия // Современные исследования социальных проблем (электрон. науч. ж-л). 2016. № 11 (67). С. 297–312.
5. Катерный И. В. Каузальные объяснения эффекта «зловещей долины» в робототехнике: теории и исследовательские данные // Качество и жизнь. 2017. № 4 (16). С. 88, 89.
6. Модели правового регулирования создания, использования и распространения роботов и систем с искусственным интеллектом / под общ. ред. В. Б. Наумова. СПб., 2019. С. 13–32, 53–67.
7. Петровский Н. А. Аналогии в юриспруденции и правоприменительной практике // Вестник Адыгейского гос. ун-та. 2005. Вып. 3. С. 128, 131.
8. Полякова Т. А., Минбале́ев А. В., Кроткова Н. В. Развитие науки информационного права и правового обеспечения информационной безопасности: формирование научной школы информационного права (прошлое и будущее) // Государство и право. 2021. № 12. С. 102, 103. DOI: 10.31857/S102694520017761-8
9. Середкина Е. В. Философские основания прикладного антропоморфизма в социальной робототехнике // Техногос. 2020. № 4. С. 56–63.
10. Столбова Н. В., Середкина Е. В., Мышкин О. С. Насколько «Зловещая долина» зловеща на самом деле? Опыт деконструкции дискурса // Вестник Пермского ун-та. Философия. Психология. Социология. 2022. Вып. 1. С. 94.
11. Филимонов В. А., Чернявская В. С. Катастрофы субъективных оценок и их модели: к вопросу инструментализации диагностики эффекта «зловещей долины» // Территория новых возможностей. Вестник Владивостокского гос. ун-та экономики и сервиса. 2021. Т. 13. № 2. С. 170–178.
12. Шевченко Л. Л. Метафора как средство моделирования концептуальной системы автора: дис. ... канд. филол. наук [Электронный ресурс]. — Режим доступа: URL: <http://www.dslib.net/germanskie-jazyki/metafora-kak-sredstvo-modelirovaniya-konceptualnoj-sistemy-avtora.html> (дата обращения: 15.06.2022).
13. Balkin J. M. The Path of Robotics Law // California Law Review Circuit. 2015. Vol. 6. P. 50, 56.
14. Belanche D., Casalo L. V., Flavián C. et al. Service robot implementation: a theoretical framework and research agenda // The Service Industries Journal. 2020. Vol. 40. Iss. 3–4. P. 208, 209.
15. Bertolini A. Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules // Law, Innovation and Technology. 2013. Vol. 5. Iss. 2. P. 217–219.
16. Calo R. Robotics and the Lessons of Cyberlaw // California Law Review. 2015. Vol. 103. No. 3. P. 529, 531.
17. Coeckelbergh M. Three Responses to Anthropomorphism in Social Robotics: Towards a Critical, Relational, and Hermeneutic Approach // International Journal of Social Robotics. 2021. P. 1 [Электронный ресурс]. — Режим доступа: URL: <https://doi.org/10.1007/s12369-021-00770-0> (дата обращения: 27.05.2022).
18. Duffy B. R. Anthropomorphism and the social robot // Robotics and Autonomous Systems. 2003. Vol. 42. P. 178, 181, 183.
19. Epley N., Waytz A., Cacioppo J. T. On Seeing Human: A Three-Factor Theory of Anthropomorphism // Psychological Review. 2007. Vol. 114. No. 4. P. 864–886.
20. Hallevy G. The Criminal Liability of Artificial Intelligence Entities — from Science Fiction to Legal Social Control // Akron Intellectual Property Journal. 2010. Vol. 4. Iss. 2. P. 174, 186–199.
21. Hu Y. Robot Criminals // University of Michigan Journal of Law Reform. 2019. Vol. 52. Iss. 2. P. 490, 492.
22. Jones C. P. A. The Robot Koseki: A Japanese Law Model for Regulating Autonomous Machines // Journal of Business and Technology Law. 2019. Vol. 14. Iss. 2. P. 413–418.
23. Kaminski M. E., Rueben M., Smart W. D. et al. Averting Robot Eyes // Maryland Law Review. 2017. Vol. 76. Iss. 4. P. 983–1025.

24. *Laakasuo M., Palomäki J., Köbis N.* Moral Uncanny Valley: A Robot's Appearance Moderates How its Decisions are Judged // *International Journal of Social Robotics*. 2021. Vol. 13. Iss. 7. P. 1679–1688.
25. *Eichocki P., Billard A., Kahn P.H.* The Ethical Landscape of Robotics // *IEEE Robotics&Automation Magazine*. 2011. Vol. 18. Iss. 1. P. 46.
26. *Lockey S., Gillespie N., Holm D., et al.* A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions // *Proceedings of the 54th Hawaii International Conference on System Sciences*. 2021. P. 5466.
27. *MacDorman K.F., Green R.D., Ho Ch.-Ch. et al.* Too real for comfort? Uncanny responses to computer generated faces // *Computers in Human Behavior*. 2009. Vol. 25. Iss. 3. P. 695–710.
28. *Mori M.* The Uncanny Valley [Электронный ресурс]. – Режим доступа: URL: <https://erikbuys.files.wordpress.com/2014/02/the-uncanny-valley.pdf> (дата обращения: 06.06.2022).
29. *Murphy R., Woods D.D.* Beyond Asimov: The Three Laws of Responsible Robotics // *IEEE Intelligent Systems*. 2009. Vol. 24. No. 4. P. 14.
30. Remembering AIBO. Nippon. Science. Technology. Feb. 2, 2017 [Электронный ресурс]. – Режим доступа: URL: <https://www.nippon.com/en/views/b00909/> (дата обращения: 02.06.2022).
31. *Richards N.M., Smart W.D.* How should the law think about robots? // In: *Robot Law* / R. Calo, A.M. Froomkin, Ia. Kerr (eds). UK, USA, 2016. P. 6, 9, 10, 16–19.
32. *Salles A., Evers K., Farisco M.* Anthropomorphism in AI // *AJOB Neuroscience*. 2020. Vol. 11. Iss. 2. P. 90, 91.
33. *Simmler M., Markwalder N.* Guilty Robots? – Rethinking the Nature of Culpability and Legal Personhood in the Age of Artificial Intelligence // *Criminal Law Forum*. 2019. Vol. 30. Iss. 1. P. 8, 20, 27.
34. *Sullins J.P.* When Is a Robot a Moral Agent? // *International Review of Information Ethics*. 2006. Vol. 6. P. 25, 28, 29.
35. *Tinwell A., Grimshaw M., Abdel Nabi D. et al.* Facial expression of emotion and perception of the Uncanny Valley in virtual characters // *Computers in Human Behavior*. 2011. Vol. 27. Iss. 2. P. 741–749.
36. *Złotowski J., Proudfoot D., Yogeewaran K. et al.* Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction // *International Journal of Social Robotics*. 2015. Vol. 7. Iss. 3. P. 353.
4. *Zilberman N.N., Stefantsova M.A.* Social robot: approaches to the definition of the concept // *Modern studies of social problems (electron. scientific journal)*. 2016. No. 11 (67). P. 297–312 (in Russ.).
5. *Kateryn I.V.* Causal explanations of the “sinister valley” effect in robotics: theories and research data // *Quality and life*. 2017. No. 4 (16). P. 88, 89 (in Russ.).
6. Models of legal regulation of the creation, use and distribution of robots and systems with artificial intelligence / under the general editorship of V.B. Naumov. SPb., 2019. P. 13–32, 53–67 (in Russ.).
7. *Petrovsky N.A.* Analogies in jurisprudence and law enforcement practice // *Herald of the Adygea State University*. 2005. Iss. 3. P. 128, 131 (in Russ.).
8. *Polyakova T.A., Minbaleev A.V., Krotkova N.V.* Development of the science of information law and legal provision of information security: formation of the scientific school of Information Law (past and future) // *State and Law*. 2021. No. 12. P. 102, 103. DOI: 10.31857/S102694520017761-8 (in Russ.).
9. *Seredkina E.V.* Philosophical foundations of applied anthropomorphism in social robotics // *Technologos*. 2020. No. 4. P. 56–63 (in Russ.).
10. *Stolbova N.V., Seredkina E.V., Myshkin O.S.* How sinister is the “Sinister Valley” really? Experience of discourse deconstruction // *Herald of the Perm University. Philosophy. Psychology. Sociology*. 2022. Iss. 1. P. 94 (in Russ.).
11. *Filimonov V.A., Chernyavskaya V.S.* Catastrophes of subjective assessments and their models: on the issue of instrumentalization of diagnostics of the “sinister valley” effect // *Territory of new opportunities. Herald of the Vladivostok State University of Economics and Service*. 2021. Vol. 13. No. 2. P. 170–178 (in Russ.).
12. *Shevchenko L.L.* Metaphor as a means of modeling the conceptual system of the author: dis. ... Candidate of Philology [Electronic resource]. – Access mode: URL: <http://www.dslib.net/germanskie-jazyki/metafora-kak-sredstvo-modelirovaniya-konceptualnoj-sistemy-avtora.html> (accessed: 06/15/2022) (in Russ.).
13. *Balkin J.M.* The Path of Robotics Law // *California Law Review Circuit*. 2015. Vol. 6. P. 50, 56.
14. *Belanche D., Casalo L.V., Flavián C. et al.* Service robot implementation: a theoretical framework and research agenda // *The Service Industries Journal*. 2020. Vol. 40. Iss. 3–4. P. 208, 209.
15. *Bertolini A.* Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules // *Law, Innovation and Technology*. 2013. Vol. 5. Iss. 2. P. 217–219.
16. *Calo R.* Robotics and the Lessons of Cyberlaw // *California Law Review*. 2015. Vol. 103. No. 3. P. 529, 531.
17. *Coeckelbergh M.* Three Responses to Anthropomorphism in Social Robotics: Towards a Critical, Relational, and Hermeneutic Approach // *International Journal of Social Robotics*. 2021. P. 1 [Electronic resource]. – Access mode: URL: <https://doi.org/10.1007/s12369-021-00770-0> (accessed: 27.05.2022).
18. *Duffy B.R.* Anthropomorphism and the social robot // *Robotics and Autonomous Systems*. 2003. Vol. 42. P. 178, 181, 183.

REFERENCES

1. *Baturin Yu. M.* Hamlet as a robot: the basics of legal regulation of artificial intelligence should be sought in the well-forgotten past of Soviet legal science: materials of the Fourth International Scientific and Practical Conference “Bachilov Readings” (February 5–6, 2021) / ed. by T.A. Polyakova, A.V. Minbaleev, V.B. Naumov. Saratov, 2022. P. 265–274 (in Russ.).
2. *Bakhteev D.V.* Artificial intelligence: ethical and legal foundations. M., 2021. P. 8–28, 118–124 (in Russ.).
3. The impact of scientific and technological progress on legal life. M., 1988. P. 114 (in Russ.).

19. *Epley N., Waytz A., Cacioppo J.T.* On Seeing Human: A Three-Factor Theory of Anthropomorphism // *Psychological Review*. 2007. Vol. 114. No. 4. P. 864–886.
20. *Hallevy G.* The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control // *Akron Intellectual Property Journal*. 2010. Vol. 4. Iss. 2. P. 174, 186–199.
21. *Hu Y.* Robot Criminals // *University of Michigan Journal of Law Reform*. 2019. Vol. 52. Iss. 2. P. 490, 492.
22. *Jones C.P.A.* The Robot Koseki: A Japanese Law Model for Regulating Autonomous Machines // *Journal of Business and Technology Law*. 2019. Vol. 14. Iss. 2. P. 413–418.
23. *Kaminski M.E., Rueben M., Smart W.D. et al.* Averting Robot Eyes // *Maryland Law Review*. 2017. Vol. 76. Iss. 4. P. 983–1025.
24. *Laakasuo M., Palomäki J., Köbis N.* Moral Uncanny Valley: A Robot's Appearance Moderates How its Decisions are Judged // *International Journal of Social Robotics*. 2021. Vol. 13. Iss. 7. P. 1679–1688.
25. *Lichocki P., Billard A., Kahn P.H.* The Ethical Landscape of Robotics // *IEEE Robotics&Automation Magazine*. 2011. Vol. 18. Iss. 1. P. 46.
26. *Lockey S., Gillespie N., Holm D. et al.* A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions // *Proceedings of the 54th Hawaii International Conference on System Sciences*. 2021. P. 5466.
27. *MacDorman K.F., Green R.D., Ho Ch.-Ch. et al.* Too real for comfort? Uncanny responses to computer generated faces // *Computers in Human Behavior*. 2009. Vol. 25. Iss. 3. P. 695–710.
28. *Mori M.* The Uncanny Valley [Electronic resource]. – Access mode: URL: <https://erikbuys.files.wordpress.com/2014/02/the-uncanny-valley.pdf> (accessed: 06.06.2022).
29. *Murphy R., Woods D.D.* Beyond Asimov: The Three Laws of Responsible Robotics // *IEEE Intelligent Systems*. 2009. Vol. 24. No. 4. P. 14.
30. Remembering AIBO. Nippon. Science. Technology. Feb. 2, 2017 [Electronic resource]. – Access mode: URL: <https://www.nippon.com/en/views/b00909/> (accessed: 02.06.2022).
31. *Richards N.M., Smart W.D.* How should the law think about robots? // In: *Robot Law / R. Calo, A.M. Froomkin, Ia. Kerr (eds)*. UK, USA, 2016. P. 6, 9, 10, 16–19.
32. *Salles A., Evers K., Farisco M.* Anthropomorphism in AI // *AJOB Neuroscience*. 2020. Vol. 11. Iss. 2. P. 90, 91.
33. *Simmmler M., Markwalder N.* Guilty Robots? – Rethinking the Nature of Culpability and Legal Personhood in the Age of Artificial Intelligence // *Criminal Law Forum*. 2019. Vol. 30. Iss. 1. P. 8, 20, 27.
34. *Sullins J.P.* When Is a Robot a Moral Agent? // *International Review of Information Ethics*. 2006. Vol. 6. P. 25, 28, 29.
35. *Tinwell A., Grimshaw M., Abdel Nabi D., et al.* Facial expression of emotion and perception of the Uncanny Valley in virtual characters // *Computers in Human Behavior*. 2011. Vol. 27. Iss. 2. P. 741–749.
36. *Złotowski J., Proudfoot D., Yogeewaran K. et al.* Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction // *International Journal of Social Robotics*. 2015. Vol. 7. Iss. 3. P. 353.

Сведения об авторах

БАТУРИН Юрий Михайлович —

член-корреспондент РАН, доктор юридических наук, профессор, заведующий кафедрой компьютерного права и информационной безопасности Высшей школы государственного аудита (факультет) Московского государственного университета им. М.В. Ломоносова; 119992 г. Москва, Ленинские горы, д. 1, стр. 13 (4-й учебный корпус)

ПОЛУБИНСКАЯ Светлана Вениаминовна —

кандидат юридических наук, доцент, ведущий научный сотрудник сектора уголовного права, уголовного процесса и криминологии Института государства и права Российской академии наук; 119019 г. Москва, ул. Знаменка, д. 10

Authors' information

BATURIN Yuri M. —

Corresponding Member of the Russian Academy of Sciences, Doctor of Law, Professor, Head of the Computer Law and Information Security Chair, Supreme State Audit School (Faculty), Lomonosov Moscow State University; 1, bld. 13 (4th academic building), Leninskie Gory, 119992 Moscow, Russia

POLUBINSKAYA Svetlana V. —

PhD in Law, associate Professor, leading researcher, Department of Criminal Law, Criminal Procedure and Criminology, Institute of State and Law of the Russian Academy of Sciences; 10 Znamenka str., 119019 Moscow, Russia