

УДК 342.7

В. В. Балановский

**Права и свободы человека и гражданина в контексте
развития и внедрения систем искусственного интеллекта**

Анализируются этический и правовой аспекты развития систем искусственного интеллекта. Этический аспект автор рассматривает с позиций философии И. Канта и описывает нравственные и социальные перспективы создания (квази)разумных роботов. Правовой аспект освещается в контексте нормативного регулирования рисков, возникающих при создании систем искусственного интеллекта. В заключение делается предположение о грядущей перестройке

правовой системы в связи с возникновением нового типа нормативного правового акта, который сочетает в себе классическую форму такого документа и машинный протокол.

Ключевые слова: Кант, искусственный интеллект, права и свободы человека и гражданина.

The author considers ethical and legal aspects of a developing of AI systems. He examines the ethical aspect through the prism of Kant's philosophy and outlines moral and social prospects of (quasi)intelligent robots. The author considers the legal aspect in context of normative regulation of risks that arise with a creating of AI systems. In conclusion the author makes an assumption on forthcoming transformation of a legal system because of a new type of legal act that combines classical form of legal act and machine protocol.

Key words: Kant, artificial intelligent, human and civil rights and freedoms.

Все большее число сфер жизни и деятельности человека подвергаются автоматизации, роботизации и цифровизации. Помимо очевидных преимуществ данный процесс сопряжен с массой сопутствующих рисков. Вызваны они, в первую очередь, коренным изменением характера коммуникативной среды, что приводит к перестройке существующих общественных отношений и, как следствие, к трансформации старых представлений о правах и свободах человека и гражданина, государстве, экономических процессах. Прогнозировать характер этих изменений довольно затруднительно, поскольку происходят они стремительно и бесконтрольно. Более того, не всегда понятно, с какой целью внедряется то или иное новшество.

Так, если взять исключительно этический аспект развития систем искусственного интеллекта (ИИ) с целью создания (квази)разумных роботов, которые способны максимально

приблизиться по своим характеристикам к человеку, станут обладать способностью к волевому инициативному нестандартному действию в нестандартных условиях, то эта цель совсем не кажется достойной. Здесь следует обратиться к И. Канту. Его знаменитая формулировка практического императива гласит:

Поступай так, чтобы ты всегда относился к человечеству и в своем лице, и в лице всякого другого также как к цели и никогда не относился бы к нему только как к средству [4, с. 205].

Практический императив адресован не только к людям, но и ко всем разумным существам. Это означает, что он может быть в полной мере распространен на (квази)разумных роботов. Следовательно, если мы и будем таковых создавать, то не сможем использовать их как средство, иначе это будет аморально. Тогда зачем тратить огромные ресурсы на конструирование бесполезного произведения чистого научно-технического творчества? Однако совершенно очевидно, что системы искусственного интеллекта, приближенные по своим параметрам к человеку, разрабатываются для того, чтобы они служили нашими инструментами, помощниками, рабами. Это означает, что хотя на определенном этапе развития человечество и пришло к мнению, согласно которому рабовладение не отвечает интересам развитого общества и представлениям о правах и свободах человека, но тем не менее продолжило предпринимать попытки вернуться к прямому (а не опосредованному) рабовладению в иной форме. В таком случае, если (квази)разумные роботы станут реальностью, то эксплуатация одних разумных существ другими снова будет нормальным и общественно приемлемым явлением. И довольно скоро может оказаться, что «стоимость» производства, поддержки и обучения «носителей естественного интеллекта» в разы ниже, чем носителей интеллекта искусственного. И тогда общество снова вернется к прямому рабовладению, тем самым откатившись в своем развитии на тысячи лет назад. Вряд ли в таких условиях что-то остановит дальнейшую деградацию человечества.

Несмотря на этические следствия идеи создания (квази)разумных машин, нужно отдавать себе отчет, что прогресс нельзя остановить. Поэтому главной задачей становится выявление и управление рисками, возникающими на этом пути.

Первый риск связан с возможностью потери контроля над искусственным интеллектом, особенно когда речь идет о машинном обучении. Сюда же относится проблема «восстания машин». Невозможно предугадать, что машина может сделать по своей квазиинициативе, к каким выводам прийти в ходе своего самообучения, поскольку объемы информации, которыми оперирует ИИ, в разы превышают человеческие возможности. Здесь же возникает фундаментальная философско-правовая задача об ответственности: кто отвечает за опасное поведение ИИ? Ведь программист, изготовитель или пользователь роботов только отчасти могут предсказать их поведение, например робота-автопилота. Возможный алгоритм предлагает Резолюция Европарламента от 16.02.2017 г. №P8_TA(2017)0051 «Нормы гражданского права в робототехнике» (далее — Резолюция) [3]. Так, в п. 55—56 Резолюции отмечается, что в случаях причинения вреда человеку устройством, управляемым ИИ, необходимо всесторонне оценивать степень автономности поведения робота, выяснять, мог ли изготовитель либо пользователь минимизировать риски опасного поведения машины в данной ситуации. При этом, как отмечается в п. 56, на современном этапе развития техники ответственность должна лежать прежде всего на человеке. В свою очередь п. 59.a и 59.b определяют в качестве главных механизмов реализации ответственности обязательное страхование и создание компенсационных фондов.

Подвидом этого риска является достижение технологической сингулярности. Это точка во времени, после которой развитие и внедрение технологий будет настолько быстрым, а сами технологии настолько сложными, что человек перестанет понимать и контролировать высокотехнологическую сферу.

На самом деле это точка, в которой прогрессирующий ИИ встретится с деградирующим естественным интеллектом. Ведь помимо угасания когнитивных способностей из-за появления умных помощников идет процесс насильственного превращения живых людей в роботов. Например, в кол-центрах операторов учат общаться с клиентами по определенном сценарию — скрипту — и только по нему. Поэтому создается впечатление, что ты говоришь не с живым человеком, а с роботом. Постепенно навязывания скриптов в общении или любой другой деятельности приводит к тому, что люди перестают творчески подходить к общению или решению задач.

Второй риск заключается в потенциальной потере контроля над теми, кто создает и эксплуатирует искусственный интеллект. Может появиться «злой гений», который запрограммирует ИИ на уничтожение человечества. Другая ситуация — когда мощный ИИ, который управляет военной инфраструктурой, может попасть в руки социопатов. В качестве одного из способов управления этим риском видится обеспечение открытости кода и процесса разработки систем ИИ. В частности, такой выход закреплен в п. 6, 9, 12, 22 Резолюции, а также в приложениях к ней — рамочном Кодексе этики разработчиков робототехники и Кодексе комитетов по этике научных исследований. Эти принципы вошли и в национальную стратегию развития ИИ во Франции [8, р. 7, 15, 114—119], представленной Э. Макроном 29 марта 2018 г. [7, р. 8].

Вместе с тем не стоит забывать, что учитывающие положения Резолюции страны ЕС не являются монополистами на рынке ИИ. Довольно мощным игроком стал Китай, где стандарты научной этики несколько отличаются от европейских, что хорошо видно на примере экспериментов по генной модификации человека [5]. В Китае нет строгих требований по открытости разработок систем ИИ, а средства выделяются колоссальные. Так, в 2017 г. Поднебесная озвучила планы по инвестированию в данную сферу 60 млрд долларов до 2025 г. Цель — безоговорочное лидерство в области ИИ. Для сравнения: Франция запланировала инвестирование в эту сферу всего 1,5 млрд евро в период с 2018 до 2022 г. [7, р. 8].

Третий риск связан с комплексом социально-экономических проблем, в первую очередь — с безработицей и несправедливым распределением благ. В той же Резолюции в п. 3 подчеркивается, что системы ИИ не должны заменять человека, а только способствовать расширению его возможностей. Правда, это сложно объяснить бизнесу, который руководствуется прагматическими принципами. Поэтому одним из выходов видится возложение дополнительных финансовых обязательств на компании ИТ-индустрии, которые, не создавая полноценных рабочих мест и не неся больших затрат, сегодня зарабатывают больше, чем гиганты традиционной экономики. И иногда возможно, что при массовой потере работы из-за внедрения ИИ система социального обеспечения позволит людям сохранить достойный уровень жизни. Однако более серьезная проблема заключается в том, что не все люди имеют привычку и способности наполнять свободное время созидательной деятельностью, что может привести к дестабилизации общества.

Четвертый риск заключается в дискриминации прав «аналоговых» людей. Это довольно свежая проблема — на наших глазах по вполне объективным причинам формируется новое меньшинство. Дело в том, что многие просто не в состоянии жить в условиях тотальной цифровизации и автоматизации. Учитывая, что стали появляться люди с физиологической аллергией на электромагнитное излучение, возникает вопрос: почему забыли о праве человека на аналоговое существование? Кто будет отстаивать интересы этой группы людей, которая не сможет или не захочет адаптироваться в мире умных машин?

Каждый из указанных рисков в той или иной степени несет угрозу закрепленным в Конституции РФ [1] правам и свободам человека и гражданина. Выход из-под контроля систем ИИ или переход контроля над ними в руки социопатов может привести к нарушению ч. 1 ст. 20; ч. 1 и 2 ст. 23; ч. 1 ст. 41; ч. 1 ст. 45 Конституции РФ. Цифровая безработица ударит по ч. 1 и 3 ст. 37. Четвертый риск затрагивает ч. 2 ст. 19; ч. 1 ст. 24; ч. 1 и 4 ст. 32; ч. 1 ст. 41; ч. 1, 2 и 4 ст. 43; ч. 1 ст. 44; ч. 1 и 2 ст. 45 Основного закона РФ.

Безусловно, нивелирование всех указанных рисков невозможно без правового регулирования. Начало этому процессу положено Резолюцией Европарламента, на основе которой пишутся национальные стратегии развития ИИ. Россия тоже не стоит в стороне от глобальных процессов. Так, распоряжением Правительства РФ от 28.07.2017 г. № 1632-р утверждена программа «Цифровая экономика Российской Федерации» на 2017—2024 гг. (далее — Программа) [2], которая предполагает в числе прочего создание правовых основ функционирования этой сферы. Однако на примере данного документа можно заметить, что на стратегическом уровне разработка и внедрение систем ИИ и формирование необходимой для этого нормативной базы несколько рассинхронизировано. В частности, мероприятие 3.3.6 Программы, предусматривающее создание системы мер финансового стимулирования и механизмов государственно-частного партнерства в сфере квантовых вычислений, ИИ, робототехники и др., запланировано на IV квартал 2018 г. В то же время мероприятие 1.8.1 по созданию благоприятных условий для решения правовых вопросов, связанных с использованием робототехники, инструментов ИИ, запланировано на II квартал 2019 г. Подчеркну: в данном пункте предполагается не создание более или менее всеохватывающей нормативной базы в 2019 г., а только формирование условий для ее появления. Таким образом, согласно Программе разработка и внедрение систем ИИ должны предшествовать нормативному регулированию данной сферы, что, учитывая более высокую скорость развития технологий по сравнению со скоростью совершенствования законодательства, может привести к неблагоприятным последствиям.

Подытоживая сказанное, отмечу, что мы наблюдаем кардинальное изменение коммуникативной среды, что неизбежно влечет изменение правовой коммуникации, а значит, если принять точку зрения А.В. Полякова, и всей правовой системы [6, с. 10]. Таким образом, вероятнее всего, назрела необходимость в новом цифровом кодексе, который бы регулировал правоотношения не только в сфере разработки систем искусственного интеллекта, но и во всех сферах, связанных с пере-

ходом к иному типу социального взаимодействия. Возможно, мы стоим на пороге появления новых субъектов (не объектов) правовой коммуникации — квазиразумных машин.

Вполне допускаю, что эти вызовы потребует кардинальной перестройки всего права, так как, например, кодекс, регламентирующий поведение ИИ, должен быть понятен не только людям, но и машинам. То есть возникнет новый вид документа — синтез свода нормативно-правовых актов и машинных протоколов. Но чтобы перевести нормативно-правовой акт в понятную машине форму, нужно, чтобы он не содержал в себе противоречий, мог быть легко переведен на язык компьютерной логики. Это повлечет за собой необходимость исключить из нормативно-правовых актов все двусмысленности и недосказанности, благодаря которым обеспечено работой огромное число юристов — от законодателей до адвокатов и судей. В итоге некий цифровой кодекс станет зародышем совершенной новой правовой системы.

Список литературы

1. Конституция Российской Федерации (принята всенародным голосованием 12.12.1993 г., с учетом поправок, внесенных законами РФ о поправках к Конституции РФ от 30.12.2008 г. № 6-ФКЗ, 30.12.2008 № 7-ФКЗ, 05.02.2014 № 2-ФКЗ, 21.07.2014 № 11-ФКЗ) // Собрание законодательства РФ. 2014. № 31. Ст. 4398.

2. *Об утверждении* программы «Цифровая экономика Российской Федерации»: распоряжение Правительства РФ от 28.07.2017 г. № 1632-р // Собрание законодательства РФ. 2017. № 32. Ст. 5138.

3. *European Parliament resolution of 16.02.2017 №P8_TA (2017) 0051 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))* // European Parliament [website]. URL: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//EN> (дата обращения: 29.11.2018).

4. *Кант И.* Основоположения метафизики нравов // Сочинения : в 8 т. Т. 4. М. : ЧОРО, 1994. С. 153—246.

5. *ГМО-дету.* Сделаны в Китае // Crdk.ru : [сайт]. URL: <https://chrchk.ru/tech/gmo-kids-born-in-china> (дата обращения: 29.11.2018).

6. Поляков А. В. Коммуникативная теория права // Коммуникативное правопонимание : избранные труды. СПб. : Алеф-Пресс, 2014.

7. Fischer S.-C. China's Rise as a New AI Superpower // ETH Globe. 2018. № 2. P. 8—9.

8. Villani C., Schoenauer M., Bonnet Y. et al. For a Meaningful Artificial Intelligence: Towards a French and European Strategy. P., 2018. URL: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf (дата обращения: 29.11.2018).

Об авторе

Балановский Валентин Валентинович, канд. филос. наук, ст. науч. сотр. Балтийского федерального университета им. И. Канта.

E-mail: v.v. balanovskiy@ya. ru

Данное исследование было поддержано из средств субсидии, выделенной на реализацию Программы повышения конкурентоспособности БФУ им. И. Канта.