МОНОГРАФИЯ

ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПРАВОВОЕ РЕГУЛИРОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, РОБОТОВ И ОБЪЕКТОВ РОБОТОТЕХНИКИ КАК УСЛОВИЕ ФОРМИРОВАНИЯ ЭКОНОМИЧЕСКОГО ЛИДЕРСТВА В РОССИИ

ИЗДАТЕЛЬСТВО Прометей

Коллектив авторов

Правовое регулирования искусственного интеллекта, роботов и объектов робототехники как условие формирования экономического лидерства в России

http://www.litres.ru/pages/biblio_book/?art=66365560
Правовое регулирования искусственного интеллекта, роботов и объектов робототехники как условие формирования экономического лидерства в России:
ISBN 978-5-00172-197-0

Аннотация

В монографии анализируется современное состояние правового регулирования искусственного интеллекта, роботов и робототехники в России и в зарубежных юрисдикциях. Проанализировано зарубежное законодательство в сфере применения киберфизических систем, искусственного интеллекта, роботов и объектов робототехники. В работе сформулирован основной понятийный аппарат в

сфере искусственного интеллекта на основе исследованных теоретических и легальных определений, сформулированы принципы развития и функционирования искусственного интеллекта. Авторами комплексно исследованы тенденции становления и развития правового регулирования искусственного интеллекта, роботов и объектов робототехники в области гражданских, социальных и экономических отношений.

Издание адресовано широкому кругу читателей: студентам, аспирантам и преподавателям юридических и технических вузов и факультетов, специалистам в сфере информационных технологий и анализа данных, а также исследователям-правоведам.

Содержание

Правовое регулирования искусственного интеллекта, роботов и объектов робототехники как условие формирования экономического лидерства в России

Авторский коллектив

Ручкина Гульнара Флюровна – профессор, декан Юридического факультета ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации», доктор юридических наук, профессор;

Демченко Максим Владимирович – кандидат юридических наук, доцент, заместитель декана Юридического факультета по научной работе ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»;

Попова Анна Владиславовна – доктор юридических наук, профессор, профессор Департамента международного и публичного права Юридического факультета ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»;

Лапина Марина Афанасьевна – доктор юридических наук, профессор, главный научный сотрудник Центра исследований и экспертиз, профессор Департамента международного и публичного права Юридического факультета ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»;

Павликов Сергей Герасимович – доктор юридических наук, профессор, руководитель Департамента правового регулирования экономической деятельности Юридического факультета ФГОБУ ВО «Финансовый университет при Прави-

тельстве Российской Федерации»;

Горохова Светлана Сергеевна – кандидат юридических наук, доцент, ведущий научный сотрудник Центра исследований и экспертиз, доцент Департамента международного и публичного права Юридического факультета ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»;

Свиридова Екатерина Александровна – кандидат юридических наук, доцент, старший научный сотрудник Центра исследований и экспертиз, доцент Департамента международного и публичного права Юридического факультета ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»;

Исмаилов Исмаил Шапурович – кандидат юридических наук, заместитель декана Юридического факультета по дополнительному профессиональному образованию и магистратуре ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»;

Баракина Елена Юрьевна – кандидат юридических наук, лаборант-исследователь Центра исследований и экспертиз, старший преподаватель Департамента международного и публичного права Юридического факультета ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации».

Введение

Государство уделяет все большее внимание необходимости регулирования новых перспективных направлений технологического развития, которые активно развиваются. Это в полной мере относится и к вопросам регулирования в области робототехники, искусственного интеллекта. В целях развития искусственного интеллекта в России указом Президента Российской Федерации от 10 октября 2019 г. № 490 утверждена Национальная стратегия развития искусственного интеллекта до 2030 года в Российской Федерации (далее – Стратегия).

Стратегией определены задачи развития искусственного интеллекта в России, к которым относятся:

– поддержка научных исследований в целях обеспечения опережающего развития искусственного интеллекта;

– разработка и развитие программного обеспечения, в котором используются технологии искусственного интеллекта;

– повышение доступности и качества данных, необходимых для развития технологий искусственного интеллекта;

– повышение доступности аппаратного обеспечения, необходимого для решения задач в области искусственного интеллекта;

– повышение уровня обеспечения российского рынка технологий искусственного интеллекта квалифицированными

кадрами и уровня информированности населения о возможных сферах использования таких технологий;

– создание комплексной системы регулирования общественных отношений, возникающих в связи с развитием и использованием технологий искусственного интеллекта.

Во исполнение Стратегии разработан федеральный проект «Искусственный интеллект» национальной программы «Цифровая экономика». В целях координации деятельности федеральных органов государственной власти, организаций бизнеса и науки постановлением Правительства Российской Федерации от 23 ноября 2019 г, № 1500 создана подкомиссия по развитию искусственного интеллекта, сопредседателями которой являются Министр экономического развития Российской Федерации и Президент, Председатель Правления ПАО Сбербанк. Главными задачами деятельности подкомиссии является рассмотрение предложений, связанных с развитием искусственного интеллекта, согласование мероприятий федерального проекта «Искусственный интеллект», содействие реализации «дорожной карты» по развитию высокотехнологичной области «Искусственный интеллект» и реализации соглашений, заключенных между Правительством Российской Федерации и ПАО Сбербанк, а также между Правительством Российской Федерации и АО «УК «РФПИ». Минэкономразвития подготовило проект концепции регулирования искусственного интеллекта (ИИ) и робототехники до 2023 г. По мнению разработ-

чиков, документ должен внести регуляторную конкретику в уже утвержденную Президентом Российской Федерации Национальную стратегию по искусственному интеллекту. Целью Концепции регулирования технологий искусственного интеллекта и робототехники является формирование основ правового регулирования новых общественных отношений, складывающихся в связи с разработкой и применением систем с искусственным интеллектом, в том числе в части создания и использования роботов, а также снятие правовых барьеров, препятствующих разработке и применению указанных систем.

Несмотря на активность государства в развитии искусственного интеллекта, роботов и объектов робототехники, многие сферы нормативного регулирования остаются не разработанными и не исследованными, особенно по сравнению со многими зарубежными странами. Это определяет актуальность проведения научных исследований, выработки теоретических и практических положений в области правового регулирования рассматриваемых в работе отношений.

В рамках исследования были определены основные задачи, решение которых обеспечивает достижение цели настоящего исследования:

– исследовать риски использования киберфизических систем, искусственного интеллекта, роботов и объектов робототехники, в том числе с учетом требований информационной безопасности;

– исследовать соотношение правовых и нравственно-этических принципов использования киберфизических систем, и разработка этических правил взаимодействия человека с искусственным интеллектом, роботами и объектами робототехники;

– провести анализ правосубъектности искусственного интеллекта, роботов и объектов робототехники;

– проанализировать возможности создания гибкой системы нормативно-правового регулирования, допускающей возможности тестирования и ограниченного использования, гарантирующая безопасность населения и направленная на стимулирование развития искусственного интеллекта, роботов и объектов робототехники;

– обосновать предложения по созданию необходимых правовых условий для эффективного использования искусственного интеллекта, роботов и объектов робототехники в сфере права интеллектуальной собственности, в банковской и транспортной сфере.

В результате исследования планируется получить научные и прикладные результаты, в частности: разработать этические правила взаимодействия человека с искусственным интеллектом, роботами и объектами робототехники; подготовить аналитический отчет о рисках использования киберфизических систем, искусственного интеллекта, роботов и объектов робототехники с учетом требований информационной безопасности; описать теоретические подходы в юрис-

пруденции к правосубъектности искусственного интеллекта, роботов и объектов робототехники и к содержанию гражданско-правовой ответственности в рассматриваемой сфере; разработать предложения по созданию необходимых правовых условий для эффективного использования искусственного интеллекта в сфере права интеллектуальной собственности, в банковской сфере, в транспортной сфере; издать монографию по материалам НИР для формирования банка рекомендуемой литературы в ходе подготовки обучающихся по соответствующим профилям программ магистратуры и аспирантуры.

Полученные результаты представлены в работе в виде конкретных выводов и рекомендаций органам государственной власти федерального и регионального уровня, к чьей компетенции относится выработка мер по созданию системы законодательного регулирования означенной сферы научного исследования в целях дальнейшего развития и упрочения цифровой экономики и достижения передовых позиций в мировом масштабе по использованию искусственного интеллекта, киберфизических систем, нейронных сетей, различных видов роботов и робототехники в Российской Федерации. Практическая значимость результатов работы состоит в возможности использования их в работе органов государственной власти и государственных органов Российской Федерации для разработки отдельных поправок, и изменений в законодательстве страны, носящих опережающий характер в

целях преодоления пробелов в сфере технологического развития страны. Теоретическая значимость исследования заключается в возможности использования результатов исследования для развития общей теории права, информационного права, гражданского права, банковского и транспортного права, права интеллектуальной собственности и иных отраслевых юридических наук; информатики, биоэтики, медицины и др. в России. Научная новизна заключается в выявлении рисков использования киберфизических систем, искусственного интеллекта, роботов и объектов робототехники, в том числе с учетом требований информационной безопасности, определения соотношения правовых и нравственно-этических принципов использования киберфизических систем, и разработке этических правил взаимодействия человека с искусственным интеллектом, роботами и объектами робототехники, установления возможности создания гибкой системы нормативно-правового регулирования, допускающей возможности тестирования и ограниченного использования, гарантирующая безопасность населения и направленная на стимулирование развития искусственного интеллекта, роботов и объектов робототехники.

Глава 1
Исследование рисков использования киберфизических систем, искусственного интеллекта, роботов и объектов робототехники, в том числе с учетом требований информационной безопасности

В настоящем разделе мы предпримем попытку рассмотрения основных вопросов, связанных с выявлением, предупреждением и минимизацией рисков в процессе использования киберфизических систем, технологий искусственного интеллекта и робототехники в различных сферах общественных отношений, а также обоснуем необходимость правового регулирования указанных процессов.

Данная тема слабо освещена в отечественной правовой научной литературе. Наиболее системно риски, связанные с использованием технологий искусственного интеллекта раскрыты в 2017 г. П.М. Морхатом в монографии «Искусственный интеллект: правовой взгляд» [1]. Имеется ряд публи-

каций и других ученых-правоведов, которые затрагивали в своих исследованиях проблемы рисков при эксплуатации искусственного интеллекта, роботов и объектов робототехники. К примеру, И.В. Понкин и А.И. Редькина в статье «Искусственный интеллект с точки зрения права» выявили имеющие существенное значение для правового регулирования основные риски и неопределенности, связанные с искусственным интеллектом [2]; в коллективной монографии под редакцией А.В. Незнамова раскрыты тенденции зарубежного законодательства в области страхования рисков при использовании роботов и технологий искусственного интеллекта [3].

Однако можно с полной определенностью говорить о взятом Россией курсе на развитие информационного общества и устойчивый социально-экономический рост в условиях поступательного изменения экономического уклада посредством интеграции определенных достижений развития цифровых технологий, о чем свидетельствует первый этап исследования по фундаментальной научной теме: «Теория правового регулирования искусственного интеллекта, роботов и объектов робототехники в Российской Федерации» [4].

Тема рисков при использовании искусственного интеллекта и объектов робототехники неоднократно поднималась в зарубежной научной литературе, причем рассмотрению предавались как широкие теоретические вопросы выявления и оценки рисков в процессе использования искусствен-

ного интеллекта и роботов в различных сферах жизнедеятельности, так и более практические и узкие разработки, которые могут быть востребованы в прикладной деятельности по созданию и эксплуатации конкретных роботизированных объектов.

Исходные понятия

Для анализа рисков использования современных роботизированных и инфокоммуникационных технологий необходимо определиться с понятийным аппаратом.

Понятия «киберфизические системы» (далее – КФС), «искусственный интеллект» (далее – ИИ) и «робот» имеют множество определений [5]. В научной правовой литературе авторский и наиболее подробный тезаурус на русском языке рассматриваемых понятий представлен П.М. Морхатом в работе «Право и искусственный интеллект: Тезаурус» [6], а также в коллективной монографии под ред. А.В. Незнамова [7].

Однако в рамках настоящей работы мы будем придерживаться следующей терминологии, которая установлена в законодательстве и в нормативно-технической документации:

а) в Федеральном законе от 24.04.2020 № 123-ФЗ и в Указе Президента РФ от 10.10.2019 № 490 **ИИ** раскрыт как комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообуче-

ние 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 (в том числе информационные системы, информационно-телекоммуникационные сети, иные технические средства обработки информации), программное обеспечение (в том числе использующие методы машинного обучения), процессы и сервисы по обработке данных и поиску решений [8, 9]; в ГОСТ Р 43.0.8—2017 и ГОСТ Р 43.0.7—2011 **ИИ** трактуется более лаконично как «моделируемая (искусственно воспроизводимая) интеллектуальная деятельность мышления человека» [10, 11];

б) в ГОСТ Р 60.0.0.4—2019/ИСО 8373:2012. Национальный стандарт Российской Федерации. Роботы и робототехнические устройства. Термины и определения» **робот**[1] (*robot*) определен весьма конкретно как «исполнительный механизм, программируемый по двум или более степеням подвижности, обладающий определенной степенью автономности и способный перемещаться во внешней среде с целью выполнения задач по назначению» [12].

---

[1] При этом в указанном ГОСТ при трактовке понятия «робот» дана ссылка на ИСО/ТК 299 «Робототехника», согласно которому в 2018 году дано новое определение: робот (robot) – программируемый исполнительный механизм с определенным уровнем автономности для выполнения перемещения, манипулирования или позиционирования.

В этом же стандарте дано понятие: «**Робототехника** (*robotics*) – наука и практика разработки, производства и применения роботов» [12]. Понятие «объект робототехники» ни в правовых, ни в технических нормах не закреплено. Исходя из приведенного понятия робототехники, к объектам робототехники можно отнести различные классы и категории роботов и робототехнические системы.

Термин **КФС** (англ. – Cyber Physical Systems, CPS) пока что нормативно не раскрыт, но, как правило, в научной литературе отечественные [14, 15, 273] и зарубежные [16, 17] специалисты в области информационной безопасности под КФС подразумевают единую информационно-технологическую среду, обеспечивающую интеграцию вычислительных ресурсов в физические процессы. Это техническая система, в которой взаимосвязаны вычислительные элементы и элементы физической природы, служащие источниками и потребителями информации.

Как видим, научная трактовка понятия КФС и нормативные, нормативно-технические трактовки понятий ИИ и робота имеют существенные различия. Так, ИИ определен на основе компьютерного моделирования весьма широко, КФС совмещает инфокоммуникационные технологии с физическими компонентами, а понятие робота не позволяет охватить широкий спектр классов и категорий роботов, которые имеют существенные различия. Исходя из приведенных терминов КФС, ИИ и роботов, можно утверждать, что подходы

к выявлению и оценке рисков их использования, в том числе с учетом требований информационной безопасности могут быть различными и обладать определенными особенностями, быть в зависимости от специфики отдельных сфер их применения.

Следовательно, общие риски с учетом обеспечения информационной безопасности, свойственные всем видам КФС, ИИ, роботам и объектам робототехники, и соответственно возможности их правового регулирования будут лежать в плоскости нормативных правовых актов в данной области. А непосредственно риски использования отдельных разновидностей КФС, ИИ и роботов должны конкретизироваться при помощи технического нормирования, поскольку необходима разработка разных правил оценки и управления рисками для различных классов устройств. Именно по этому пути пошли как страны ЕС в развитии европейской роботизированной промышленности, исходя из смысла содержания Резолюции Европейского Парламента (далее – Резолюция) [18], так и наше государство.

Так, в Резолюции содержатся рекомендации по установлению критериев классификации роботов; по необходимости прохождения специальной регистрации и об обязательности оснащения роботов «черным ящиком», в котором записываются данные по каждой выполняемой операции, включая логику принятия решений; обоснование разработки правил тестирования новых роботов в реальных условиях; обосно-

вание необходимости внедрения обязательного страхования для определенных видов роботов.

Важным документом для исследования рисков использования технологий ИИ является Руководство по этике для ИИ, принятое в ЕС [18]. Одним из трех основных принципов использования ИИ является принцип «не навредить». Руководством установлены общие правила разработки таких технологий: они должны быть безопасными, подотчётными, не носить дискриминационного характера, быть поднадзорными человеку.

Перечисленные документы ЕС, по сути, являются базовым юридическим фундаментом правового регулирования риск-ориентированного подхода использования современных роботизированных и инфокоммуникационных технологий.

В Российской Федерации проблемы обеспечения информационной безопасности при использовании КФС, ИИ, роботов и объектов робототехники путем учета, оценки и управления рисками находит свое отражение как в нормативных актах различного уровня (законы, указы Президента, постановления Правительства РФ, ведомственные нормативно-правовые акты), так и в нормативно-технических документах (технические регламенты, национальные стандарты).

Безусловно, что для успешного функционирования роботизированных и инфокоммуникационных технологий долж-

ны учитываться возникающие в результате такой деятельности риски [19].

Что такое риски?

Этимология слова «риск» (из французского *risque* – «риск», итальянского *risico*) *произошло из* др. – греч. ῥιζικόν — «утёс», др. – греч. ῥίζα – «подножие горы»; «рискова́ть» – первоначально означало «лавировать между скал» [20]. Классически правовой риск рассматривается как присущая человеческой деятельности объективно существующая и в определённых пределах способная к оценке и волевому регулированию вероятность понесения субъектами правоотношений негативных последствий вследствие наступления неблагоприятных событий, закономерно связанных с разнообразными предпосылками (факторами риска) [21].

Следует также отметить активно развивающуюся реляционную теорию риска. Boholm Å., Corvellec H., раскрывая ее сущность, отмечают, что указанная теория выводит риски из расположенного познания, которое устанавливает отношения риска в контексте определенных непредвиденных обстоятельствах и определенном причинном воздействии. Сильная сторона реляционной теории, согласно концепции авторов, состоит в том, что она позволяет интерпретировать природу риска; отвечать вопросы о том, что относится к рискам и почему, а также предлагает новые подходы к управлению

рисками, так как повышает уровень информированности о них всех заинтересованных субъектов [22].

Так, Head G.L.Э рассматривая теоретические вопросы корпоративного управления всевозможными рисками, делает акцент на субъективные факторы продуцирования рисков. Разные поколения управленцев, находившихся у руля управления корпорациями, имеют разные взгляды на один и тот же риск. Это касается также отдельно взятого индивида. Субъективные взгляды управленцев на корпоративном уровне могут иметь решающее значение на организацию управления реалиями рисков. Меняющаяся стратегия управления рисками может резко изменить действия менеджеров по управлению рисками [23].

Риск в классическом понимании в любых сферах жизнедеятельности людей (без использования КФС, ИИ или роботов) носит двойственный субъект-объектный характер, соответственно, элементы риска подразделяются на объективные (факторы и ситуация риска) и субъективные (субъект и волевое регулирование)» [21, с. 20]. Как правило, в гражданско-правовых отношениях «правовой риск – это текущий или будущий риск потери дохода, капитала или возникновения убытков в связи с нарушениями или несоответствием внутренним и внешним правовым нормам, таким как законы, подзаконные акты регуляторов, правила, регламенты, предписания, учредительные документы» [21, с. 21].

Представляется важным принятая в национальных стан-

дартах России терминология рисков. Пункт 1.1 Национального стандарта Российской Федерации ГОСТ Р 51897—2011/Руководство ИСО 73:2009 от 16.11.2011 «Менеджмент риска. Термины и определения» [24] и пункт 3.1 Национального стандарта Российской Федерации ГОСТ Р ИСО 31000—2019 «Менеджмент риска. Принципы и руководство» [25], дают определение понятию риска как «***следствия влияния неопределенности на достижение поставленных целей***», давая также нижеследующие разъяснения: «Под следствием влияния неопределенности необходимо понимать отклонение от ожидаемого результата или события (позитивное и/ или негативное). Цели могут быть различными по содержанию… и назначению (стратегические, общеорганизационные, относящиеся к разработке проекта, конкретной продукции и процессу). Риск часто характеризуют путем описания возможного события и его последствий или их сочетания. Риск часто представляют в виде последствий возможного события (включая изменения обстоятельств) и соответствующей вероятности. Неопределенность – это состояние полного или частичного отсутствия информации, необходимой для понимания события, его последствий и их вероятностей».

Следует отметить, что в научной литературе высказывается точка зрения о том, что восприятие дефиниции риска в контексте неопределенности не является правильным. Так, T Ward S., Chapman C. разграничивают понятия «неопреде-

ленность» и «риск», отмечая при этом, что термин «риск» в большей степени поощряет перспективу угрозы, а неопределённость имеет широкое смысловое толкование. Акцент на неопределенности, по мысли авторов, улучшит управляемость рисками при реализации проекта, так как предоставит дополнительный управленческий ресурс для управления возможностями [26].

Статьей 2 Федерального закона от 27.12.2002 № 184-ФЗ (действ. ред.) «О техническом регулировании» понятие риск определяется как «вероятность причинения вреда жизни или здоровью граждан, имуществу физических или юридических лиц, государственному или муниципальному имуществу, окружающей среде, жизни или здоровью животных и растений с учетом тяжести этого вреда» [27]. Понятие «риск» неразрывно связано с понятием «безопасность». Авторы комментария к закону подчеркивают, что «даже после принятия всех мер безопасности некоторый риск, именуемый, как правило, остаточный, всегда будет присутствовать. Такой риск рассматривается как допустимый, т. е. приемлемый для каждой конкретной ситуации с учетом существующих общественных ценностей (в т. ч. экономических, политических факторов, традиций) в конкретной стране и в конкретное время» [28]. Безусловно, при оценке и управлении рисками при использовании КФС, ИИ, роботов и объектов робототехники принцип допустимости рисков будет лежать в основе правового регулирования.

Нормативно определенные риски использования информационных и инфокоммуникационных технологий

Практически любая отрасль российского законодательства регулирует в той или иной мере отношения, связанные с техническим прогрессом, включая использование современных информационных и инфокоммуникационных технологий, являющихся «фундаментом» технологий искусственного интеллекта или готовых роботизированных изделий. Административное, информационное, финансовое, таможенное, экологическое законодательство регламентируют происходящие в обществе процессы цифровизации, нормативно придают им общественно-значимый и социально-ориентированный характер. В свою очередь риски, связанные с использованием КФС, ИИ и др. роботизированных технологий как значимое явление общественной жизни напрямую воздействуют не только на социально- экономические отношения, но и политические, экологические, а также и юридические отношения. К примеру, риски использования информационных и инфокоммуникационных технологий в финансовой сфере, возникая в недрах экономических отношений, оказывают влияние на правовую, политическую и финансовую системы страны. Риски использования информационных и инфокоммуникационных технологий в реальных сек-

торах экономики (промышленности – особо опасные производственные объектов, электроэнергетике, недропользовании, сельском хозяйстве и др.) оказывают влияние на социальную сферу, экологические отношения, в конечном итоге правовую сферу.

Нельзя не отметить и обратную тенденцию. Риски, генерированные правовыми актами в области использования информационных и инфокоммуникационных технологий, порождают экономические, социальные, экологические и политические деформации.

В Российской Федерации в целом ряде законов и подзаконных нормативных правовых актов содержатся нормы, предусматривающие анализ, оценку, прогнозирование и минимизацию рисков в различных сферах общественных отношений.

При этом базовым нормативным правовым актом в части государственного управления рисками в Российской Федерации является Федеральный закон от 28 июня 2014 г. № 172-ФЗ «О стратегическом планировании в Российской Федерации» [29]. Данный закон регулирует отношения, возникающие между участниками стратегического планирования в процессе целеполагания, прогнозирования, планирования и программирования социально-экономического развития Российской Федерации, субъектов Российской Федерации и муниципальных образований, отраслей экономики и сфер государственного и муниципального управления,

обеспечения национальной безопасности Российской Федерации, а также мониторинга и контроля реализации документов стратегического планирования.

Согласно пункту 5 статьи 3 указанного Федерального закона, «прогнозирование – деятельность участников стратегического планирования по разработке научно обоснованных представлений **о рисках социально-экономического развития, об угрозах национальной безопасности Российской Федерации**, о направлениях, результатах и показателях социально-экономического развития Российской Федерации, субъектов Российской Федерации и муниципальных образований». В пункте 21 статьи 3 стратегический прогноз Российской Федерации понимается как «документ стратегического планирования, содержащий систему научно обоснованных представлений **о стратегических рисках социально-экономического развития и об угрозах национальной безопасности** Российской Федерации». Статья 7 вышеуказанного Федерального закона в числе принципов стратегического планирования выделяет принцип реалистичности, смысл которого определяется следующим образом: «при определении целей и задач социально-экономического развития и обеспечения национальной безопасности Российской Федерации участники стратегического планирования должны исходить из возможности достижения целей и решения задач в установленные сроки **с учетом ресурсных ограничений и рисков**» (пункт 9

статьи 7). Статья 16 определение **рисков** при достижении целей социально-экономического развития Российской Федерации и целевых показателей на долгосрочный период, с учетом задач обеспечения национальной безопасности Российской Федерации, установлено в качестве требования к подлежащей к разработке каждые 6 лет Стратегии социально-экономического развития Российской Федерации (подпункт 2 пункта 7 статьи 16). Стратегический прогноз Российской Федерации, разрабатываемый по поручению Президента Российской Федерации на двенадцать и более лет Правительством Российской Федерации, также должен включать «оценку рисков социально-экономического развития и угроз национальной безопасности Российской Федерации» (подпункт 1 пункта 2 статьи 23), разработку «оптимального сценария преодоления рисков и угроз с учетом решения задач национальной безопасности Российской Федерации» (подпункт 3 пункта 2 статьи 23). В числе основных задач мониторинга реализации документов стратегического планирования заявлены «проведение анализа, выявление возможных рисков и угроз и своевременное принятие мер по их предотвращению» (подпункт 6 пункта 2 статьи 40) [29].

Помимо вышеозначенного Федерального закона, нормы, предусматривающие необходимость управления рисками, встречаются, хотя и нечасто, в других нормативных правовых актах [30, 31, 32, 33, 34].

На необходимость учета рисков в различных сферах неод-

нократно обращал внимание Президент Российской Федерации в своих ежегодных посланиях. Так, в Послании Федеральному Собранию от 20.02.2019 Президент РФ подчеркнул, что «работать на стратегические цели необходимо уже сегодня… Мы обязаны двигаться только вперед, постоянно набирая темп этого движения» [35]. В монографии А.О. Турганбаева при исследовании перспектив развития административно-правового обеспечения стратегического планирования в государственном управлении сделан вывод о технологическом отставании России по внедрению новейших технологий в стратегическое планирование от государств как англосаксонской правовой системы (Австралия, Великобритания, Канада, США), так и некоторых государств Азиатско-Тихоокеанского региона (Китай, Сингапур, Таиланд, Южная Корея, Япония) [36, с. 144]. Только благодаря задействованию современных информационных и инфокоммуникационных технологий и систем искусственного интеллекта в стратегическом планировании, возможно, достичь главной цели государства – улучшение условий жизнедеятельности людей и повышение комфорта жизненной среды. Однако предвидеть многие риски, вызовы и проблемы, возникающие с использованием этих технологий будущего необходимо уже сейчас.

Риски использования киберфизических систем (КФС)

Отечественными учеными Куприяновским В.П., Намиотом Д.Е., Синяговым С.А. отмечено, что современные КФС «интегрирует в себе кибернетическое начало, компьютерные аппаратные и программные технологии, качественно новые исполнительные механизмы, встроенные в окружающую их среду и способные воспринимать ее изменения, реагировать на них, самообучаться и адаптироваться» [37, с. 22]. При этом неотъемлемым свойством КФС является связанность их физических компонентов посредством инфокоммуникационных технологий [38]. КФС соединяет физические процессы производства или иные другие процессы программно-электронными системами (к примеру, систему управления распределения электроэнергии), реализуемые путем непрерывного управления [39]. Это является отличительной чертой КФС и в то же время их слабым местом. Возможность удалённого доступа к физическим компонентам (оборудованию, автомобилям, кардиостимуляторам) даёт злоумышленнику возможность перехватить управление над ними. Таким образом, возникает риск, связанный с возможностью несанкционированного доступа, перехвата и злоумышленного изменения процесса управления физическим компонентом.

К примеру, в 2009 году вредоносное программное обеспечение Stuxnet вывело из строя центрифуги на иранском заводе по обогащению урана. В результате атаки иранская атомная промышленность была отброшена на несколько лет назад. В 2019 году атаке подверглась энергетическая система Венесуэлы. Злоумышленники получили контроль над системой управления электроснабжением в столице Венесуэлы и над системой управления гидроэлектростанции имени Симона Боливара и дистанционно отключили электрическую сеть. В результате атаки половина страны осталась без света.

Риск проникновения злоумышленника в киберфизическую систему не ограничивается одним ущербом экономике и промышленности государства. Последствием атаки на КФС может стать и смерть человека. Так, в проведённых в лабораторных условиях экспериментах исследователи продемонстрировали возможность удалённого доступа к кардиостимуляторам и изменения режима их работы, в результате чего человек, которому они имплантированы, может умереть. Также проводились эксперименты по получению удалённого доступа к системе управления автомобилем. Марин Ивезич, эксперт по информационной безопасности, указывают на трудность обнаружения следов таких атак. По мнению специалистов, криминалисты, расследующие подобные инциденты, «скорее всего, не обратят внимание на немногочисленные оставленные следы и сочтут смерть случайной» [40].

Итак, первый критерий классификации рисков использования КФС можно сформулировать как риск нанесения ущерба от несанкционированного доступа к удалённым устройствам (физическим компонентам) и их потенциальной злоумышленной компрометации.

Вторым критерием классификации рисков зачастую выступают три известных свойства информации: конфиденциальность, целостность, доступность (в английской номенклатуре – «CIA»: Confidentiality, Integrity, Availability).

По критерию CIA риски классифицируются в зависимости от влияния на вышеперечисленные свойства информации:

• Риск нарушения конфиденциальности информации в киберфизических системах;

• Риск нарушения целостности информации в киберфизических системах;

• Риск нарушения доступности информации в киберфизических системах.

При этом не без основания подчёркивается, что в киберфизических системах наибольшую опасность представляют риски целостности и доступности информации [41, 42]. Например, если злоумышленник получит доступ к информации, которая содержится в кардиостимуляторе, он сможет узнать ритм работы сердца, о возможных нарушениях в его работе, повседневном графике человека. Эта информация хоть и является персональными данными и охраняется зако-

ном, но нарушение её разглашение не так опасно, как нарушение целостности и доступности информации в кардиостимуляторе, которая может привести к сбоям в его работе или вообще к выводу прибора из строя, что для человека может иметь весьма трагичные последствия. Классификация рисков по свойствам информации является достаточно условной, так как многие угрозы воздействуют сразу на несколько вышеперечисленных свойств информации.

К тому же эксперты в области информационной безопасности сходятся во мнении, что общепринятая классификация рисков по свойствам информации не является исчерпывающей при использовании КФС. Так, Хью Бойз, руководитель отдела по кибербезопасности Института техники и технологий (Institution of Engineering and Technology), выделяет ещё два значимых для КФС свойства: управляемость/контроль и полезность [43]. В случае нарушения управляемости оператор системы, хотя может и видеть наличие проблемы, но исправить ситуацию будет уже не способен. В случае нарушения контроля за системой, оператор, даже имея возможность воздействовать на систему, не будет получать корректные сведения о её состоянии. Обращая внимания на полезность, как на свойство ХВС, Хью Бойз, приводит в пример потерю космического аппарата, целью которого был сбор климатической информации о Марсе. В ходе его разработки одна проектная группа внедрила систему, использующую метрические единицы измерения (км/ч), а другая груп-

па использовала имперские единицы измерения (миль/час). В результате аппарат неправильно вошёл в атмосферу Марса и был уничтожен.

Вышеописанные риски с уверенностью могут быть применены к роботам и объектам робототехники, ведь роботы являются разновидностью КФС.

Риски использования роботов и объектов робототехники

Несмотря на то, что роботы хоть и могут действовать в определённой степени автономно, им всё же необходимы каналы связи с оператором. Возможность действовать полностью автономно у роботов может появиться только после внедрения в них технологий ИИ. Но этот шаг может принести даже больше рисков от использования роботов, чем имеется в настоящий момент. Однако, перейдём к рассмотрению роботов, как разновидности КФС.

Самой распространённой классификацией рисков использования роботов, соответственно, является вышеописанная при использовании КФС **классификация рисков по свойствам информации**: конфиденциальности, целостности и доступности. Анализируя разновидности рисков, эксперты по информационной безопасности акцентируют внимание на возможности утраты конфиденциальности бесед и тайны частной жизни в результате использования роботов. Хоть

проблема утечки информации не является самой опасной для промышленных роботов и кардиостимуляторов, она является достаточно распространённой для роботов, использующихся в бытовых целях. В условиях домашнего использования роботов, когда отказ машины не приведёт ни к чему критичному, кроме чувства впустую потраченных денег, именно риск нарушения тайны частной жизни может иметь наибольшую опасность. Так, президент фонда «prpl Foundation» Арт Свифт утверждает, что производители роботов, спеша выйти на рынок со своей моделью, зачастую забывают об обеспечении безопасности своих изделий, результатом чего может стать утечка информации [44].

Не гораздо дальше в плане обеспечения безопасности ушли и производители промышленных роботов. В отчёте Trend Micro, компании, специализирующейся на обеспечении информационной безопасности, приведены результаты исследования роботов, предназначенных для промышленного производства от ряда известных производителей на предмет их информационной безопасности [45]. В результате исследования специалисты пришли к выводу, что все испытуемые роботы уязвимы к атаке извне. По мнению Дэна Вебера, технического директора компании Mocana: «Один из самых тревожных выводов в докладе Trend Micro об уязвимостях роботов, используемых в производстве – это то, как легко хакерам, в данном случае “исследователям”, обнаружить незащищённые промышленные устройства онлайн» [46]. Тре-

вожно это потому, что злоумышленникам не составит труда не только проникнуть в сеть предприятия, но даже найти уязвимые точки для проникновения.

Кроме классификации рисков по свойствам информации существует также **классификация рисков по видам атак**, которые возможны в отношении роботов:

• *Риск атаки, модифицирующей намерение*. Это атака, направленная на искажение сообщений, передающихся к роботу. Она может преследовать цель, как изменения поведения робота, так и вывода его из строя. К данному типу атак так же относится и, так называемая, атака «отказа в обслуживании» (Denial of Service). Смысл данной атаки заключается в том, что робот оказывается перегружен входным трафиком. В результате чего происходит или остановка его работы, или задержка в реакции на поступающие команды

• *Риск атаки, манипулирующей намерением.* Это атака, направленная на модификацию сообщений, передающихся от роботов. Цель данной атаки очевидна – исказить сведения о состоянии машины.

• *Риск захватывающей атаки.* Это атака, в ходе которой злоумышленник полностью берёт под контроль связь между роботом и его оператором [47].

Вышеописанные классификации рисков сформированы исключительно через призму информационной безопасности. И это оправдано для КФС и роботов, которые согласно приведенным в начале раздела данного отчета понятий-

ям являются просто машинами (или совокупностью машин) и представляют опасность только в случае их неисправной работы, вызванной внутренним сбоем или вмешательством извне.

Такие риски будут являться специфическими, и они могут зависеть от конкретного роботизированного устройства или применяемой инфокоммуникационной технологий в КФС.

В Российской Федерации принят ряд национальных стандартов, устанавливающих требования по безопасности эксплуатации роботов с учетом оценки рисков их использования. К примеру, в «ГОСТ Р 60.1.2.1—2016/ ИСО 10218—1:2011. Национальный стандарт Российской Федерации. Роботы и робототехнические устройства. Требования по безопасности для промышленных роботов. Часть 1. Роботы» [48] в приложении А в форме таблицы приведен список существенных опасностей для робота и его подсистем, состоящий из 10 типов или групп опасностей (механические, электрические, термические, эргономические; опасности от шума, вибраций, излучения, материалов/веществ; опасности, связанные с внешней средой, в которой используется машина; комбинации опасностей). Чтобы идентифицировать любые опасности, которые могут возникнуть, должен быть осуществлен анализ опасностей.

Общая оценка рисков должна быть выполнена для всех опасностей, выявленных при идентификации опасностей.

В «ГОСТ Р 60.2.2.1—2016/ИСО 13482:2014. Национальный стандарт Российской Федерации. Роботы и робототехнические устройства. Требования по безопасности для роботов по персональному уходу» список существенных опасностей роботов, предназначенных для персонального ухода, содержит 85 разновидностей. Данным стандартом предусмотрена общая оценка рисков, включающая не только идентификацию опасности с целью выявления любых опасностей, которые могут возникнуть для конкретного робота по персональному уходу, но и оценку рисков. При этом оценка риска должна быть выполнена для всех опасностей, выявленных при идентификации опасностей, «с обращением особого внимания на разные ситуации, в которых робот по персональному уходу может контактировать с объектами, связанными с безопасностью.

После того как все меры по безопасности конструкции и защите приняты, должен быть оценен остаточный риск робота по персональному уходу и должно быть обосновано, что этот риск снижен до приемлемого уровня» [49].

Благодаря системе технического нормирования [50] можно своевременно и гибко регулировать со стороны государства риски использования роботов и объектов робототехники.

Риски использования искусственного интеллекта (ИИ)

ИИ представляет собой совершенно новую сущность, субстанцию, одно только использование которого уже рождает множество этических, правовых и технических проблем. В отличие от роботов и киберфизических систем ИИ способен самостоятельно принимать решения и самообучаться. Негативные последствия от функционирования ИИ могут быть результатом не только откровенной ошибки при его разработке, иногда такие последствия становятся результатом самостоятельных действий ИИ, разработчики которого даже не могли их предугадать. В связи с этим использование ИИ само по себе может нести определённые угрозы, которые необходимо учитывать.

Как правило, риски использования ИИ подразделяют на общие и применительно к сфере информационной безопасности.

В подготовленном исследователями Оксфордского университета докладе под редакцией Пьерлуиджи Поганини, посвящённом основным угрозам человечеству, ИИ входит в число 12 рисков, которые потенциально могут уничтожить человеческую расу [51].

Данные риски в докладе классифицированы по 4 группам: Текущие риски; Внешние риски; Развивающиеся риски; Рис-

ки глобальной политики.

К первым исследователи отнесли глобальное потепление и угрозу ядерной войны. Ко вторым – возможность столкновения Земли с астероидом. Под рисками глобальной политики исследователи понимают риски, связанные с глобальным правительством. К развивающимся рискам были отнесены риски, создающиеся руками человека, как, например, синтетическая биология и ИИ. Относительно ИИ в докладе поддерживается следующее мнение: «Искусственный интеллект, похоже, обладает огромным потенциалом для целенаправленной работы по уничтожению человеческой расы. Хотя, синтетическая биология и нанотехнологии наряду с искусственным интеллектом могут стать ответом на многие существующие проблемы, однако, при неверном использовании, это, вероятно, может быть худшим инструментом против человечества» [51].

Анализируя риски использования ИИ, следует четко разграничивать понятия «сильного искусственного интеллекта» и «слабого искусственного интеллекта». Сильный ИИ – это интеллект, который подобно человеческому, может решать различные задачи, мыслить, адаптироваться к новым условиям [52]. То есть, по сути, интеллект, способный выполнять все те же функции, которые выполняет интеллект человека. На данный момент такого интеллекта не существует. Существует лишь слабый ИИ – интеллект, способный выполнять узкоспециализированные задачи [52].

Несмотря на то, что сильный ИИ – вещь из разряда фантастики, многие исследователи склонны считать, что его разработка вполне возможна. По мнению экспертов «Центра изучения экзистенциальной угрозы» при Кембриджском университете создание «сверхразума» возможно уже в этом столетии [53]. В связи с этим эксперты Центра отмечают наиболее высокие риски, связанные с таким ИИ. Они их разделяют на риски, связанные с несчастными случаями (safety risks), и на риски, связанные со злоупотреблением таким ИИ (security risks). К первым они относят возможность выхода из строя ИИ со всеми вытекающими катастрофическими последствиями (особенно, в случае если от функционирования ИИ зависит работа критически важной инфраструктуры). Ко вторым отнесены угроза попадания технологий в руки «плохих актёров» на международной арене и угроза дестабилизирующей гонки вооружений в области ИИ[2]. О рисках внедрения сильного ИИ можно много рассуждать, но это имеет мало практического смысла, так как, о чём уже говорилось выше, данные технологии отсутствуют в настоящее время. Поэтому перейдём к рискам использования слабого ИИ.

Так, эксперты говорят о возможности манипуляции общественным мнением посредством искусственного интеллекта [54]. Медиа-ресурсы уже давно и достаточно успешно используют автономные алгоритмы для целевого маркетинга. Но, если ИИ умеет подбирать интересующие пользователя

[2] Аббревиатура ИИ, обозначает искусственный интеллект; РТ – робототехника.

товары, он также может, используя определённые личные данные, представить ему необходимую информацию в том виде и в том формате, в котором он сочтёт её наиболее достоверной, тем самым манипулируя восприятием пользователя.

Риск манипулирования общественным мнением вытекает из риска вторжения в личную жизнь пользователя посредством ИИ. Жизнь современного человека и так находится под постоянным присмотром различных систем сбора и анализа данных, начиная с сервисов целевой рекламы, и заканчивая камерами видеонаблюдения, которыми оборудованы все крупные мегаполисы. ИИ, в свою очередь, является эффективным инструментом для анализа всей собираемой информации. Аккумулируя информацию из нескольких источников, искусственный интеллект может достаточно точно формировать психологический и поведенческий портрет человека. Определять сферу его интересов, круг общения и многое другое. Сбор подобной информации о человеке может принести не только внутренний дискомфорт, но и определённые негативные материальные последствия. Например, в Китае уже введена система социального кредитования, согласно которой каждому гражданину присвоен личный балл на основании его поведения. Система оценивает надёжность граждан Китая по разным критериям, в том числе, по тому, ходят ли они по улице, покупают ли китайские товары, что они размещают в интернете. Люди с высокими баллами

получают право получить скидки на счета по электроэнергии или лучшие процентные ставки по вкладам. Граждане с низким рейтингом наоборот ограничиваются в правах. Например, по некоторым данным, им может быть ограничен высокоскоростной интернет. По сведениям CBS почти 11 миллионов китайцев из-за введения рейтинга кредитования не смогут воспользоваться услугами авиаперевозчиков, а 4 миллиона – поездами [55]. В связи с этим отмечается такой риск использования ИИ как **риск дискриминации**.

Также эксперты отмечают ещё и **риск несогласованности целей машин и людей**. Связан он, тем, что команды, в которые человек закладывает определённый смысл, могут быть совершенно по-иному интерпретированы машиной. Например, команда «доставь меня в аэропорт как можно быстрее» может иметь крайне негативные последствия. Если не уточнить, что нужно соблюдать правила дорожного движения, так как человеческая жизнь ценнее потерянного времени, то ИИ может дословно исполнить указание и оставить за собой шлейф несчастных случаев и аварий.

Наносить вред ИИ может не только в результате недоработок разработчиков и неверной интерпретации команд, но и в результате определённого внешнего воздействия. Для того чтобы оценить опасность такого негативного влияния извне на ИИ обратимся к рискам использования ИИ с позиции информационной безопасности.

Но прежде следует сказать, что ИИ – это инструмент,

средство достижения тех или иных целей и, соответственно, он, как и любой другой инструмент, может использоваться как в общественно полезных целях, так и в преступных целях. Специалисты в области информационной безопасности ожидают, что уже в ближайшем будущем ИИ может появиться на вооружении киберпреступников. Об этом, в частности, говорит Дэвид Капуано, коммерческий директор компании BluVector, занимающейся разработками решений в области кибербезопасности [56]. Эксперт предупреждает об опасности проведения злоумышленниками более мощных и неуловимых атак с использованием технологий ИИ. По мнению Д. Капуано, несмотря на то, что сейчас технологии осуществления атак с использованием ИИ находятся в зачаточном состоянии, в ближайшее время они будут стремительно развиваться. Способствовать этому будет увеличение потенциальной прибыли в результате проведения подобных атак и то, что между киберпреступниками налажен более активный обмен информацией и инновациями, чем между теми, кто им противостоит.

Эксперты предполагают, что ИИ может найти применение при осуществлении преступниками фишинговых атак (то есть, атак, направленных на обман пользователя, в результате чего он либо передаёт конфиденциальную информацию злоумышленнику, либо загружает вредоносное программное обеспечение на своё устройство). ИИ может способствовать автоматизации таких атак или упростить их про-

ведение путём, например, синтезирования голоса человека [57]. Злоумышленники могут использовать ИИ также для поиска уязвимостей и для автоматизированного использования обнаруженных уязвимостей. Стоит подчеркнуть, что и сейчас существуют инструменты, автоматизирующие вышеуказанные процессы, однако, в случае использования ИИ, процесс поиска уязвимостей и их эксплуатации будет происходить совершенно в других масштабах и на совершенно ином уровне. В случае злонамеренного использования технологий ИИ возможно распространение ботнетов (компьютерных сетей, состоящих из устройств, зараженных вредоносным программным обеспечением, позволяющим использовать данные устройства в своих целях) без единого сервера управления. Борьба с такими ботнетами значительно усложнится, ведь обычно используемая тактика отключения управляющего сервера по отношению к таким ботнетам будет неэффективна. ИИ может быть использован для координации работы в автоматическом режиме обычных ботнетов, позволяя значительно расширить масштаб хакерских атак, в которых задействуются данные ботнеты. По мнению экспертов в области информационной безопасности серьёзную опасность представляет возможность создания искусственным интеллектом новых образцов вредоносного программного обеспечения. ИИ, анализируя недостатки вредоносного программного обеспечения, созданного человеком, способен сгенерировать более продвинутые его формы, значи-

тельно затрудняя обнаружение и нейтрализацию таких программ [58]. Кроме того, используя ИИ для разработки вредоносного программного обеспечения, киберпреступники могут сделать его устойчивым к сканированию тем искусственным интеллектом, задачей которого является обнаружением вредоносных программ [59].

Опасность представляет не только возможность использования технологий ИИ злоумышленниками, но и вполне благовидное использование программного обеспечения с поддержкой ИИ. По мнению Алекса Смита, директора по решениям в области информационной безопасности компании Intermedia, в результате распространения продуктов, использующих ИИ, компании столкнутся с ростом числа случайных уязвимостей и нарушений информационных безопасности, не связанных с действиями хакеров. В частности, эксперт отмечает, что без защиты могут остаться наборы данных, на основе которых обучался ИИ. Разглашение такой информации может иметь негативные последствия, как для имиджа компании, так и для тех, чьи сведения содержатся в такой выборке (особенно, если выборки составлены на основе данных из социальных сетей).

Кроме того, программные решения, основанные на технологиях ИИ, имеют свои слабые места и уязвимости. Например, исследователи отмечают следующие наиболее высокие риски по отношению к подобным программным решениям в результате эксплуатации тех или иных уязвимостей в алго-

ритме искусственного интеллекта [60]:

1. Риск атаки путём представления специально сформированных входных данных, чтобы исказить работу алгоритма искусственного интеллекта. Суть данной атаки заключается в том, что атакующий, вычисляя ошибки на выходе алгоритма по отношению к входным данным, может таким образом сформировать входные данные, чтобы на выходе получить интересующий его результат. Например, сформировав определённым образом код вредоносного программного обеспечения, атакующий может добиться того, чтобы проверяющий этот код искусственный интеллект счёл его заслуживающим доверия. Данная атака возможна в силу того, что выборка данных, на основании которых алгоритм проходил обучение, не всеобъемлюща. При получении входных данных, которые не содержались в обучающей выборке, алгоритм может выдать случайный результат.

2. Риск нарушения конфиденциальности обучающих данных. Атакующий, наблюдая, как алгоритм реагирует на те или иные входные данные, может вычислить те данные, на основании которых алгоритм обучался.

Конец ознакомительного фрагмента.