

Миронова Наталия Геннадьевна

БЕЗОПАСНОСТЬ ИСПОЛЬЗОВАНИЯ КОГНИТИВНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ПРИНЯТИЯ РЕШЕНИЙ

Аннотация: в работе рассматривается интеллектуальная автоматизация процедур принятия решений и управления, которая внедряется во многие области социально-экономической практики, включая финансово-кредитные бизнес-процессы, в торговлю и электронную коммерцию (профилирование клиентов, маркетинговое микро-таргетирование), телекоммуникации, промышленность (технологический контроль, робототехника, нейроуправление, стратегическое планирование и прогнозирование), а также интеллектуальная автоматизация пришла в бизнес-управление, в государственное управление. Утверждается, что расширяется автоматизация управления персоналом (контроль за соблюдением требований, профилирование и оценка KPI, прогнозирование конфликтов и нарушений), медицине, транспортной отрасли и сельском хозяйстве используется беспилотная техника и другая нейросетевая автоматика. Интеллектуальные технологии приходят в образование (так, в Москве апробируется система предиктивной аналитики цифрового следа учащихся для оптимизации и таргетирования образовательных услуг, помощи ориентации в будущей профессии). Использование когнитивных технологий при создании экспертных, советующих систем, систем поддержки принятия решений дает не только удобства и экономию времени и усилий, но порождает разнообразные организационные, экономические, этические, социальные проблемы, порождая новые риски. В исследовании выполнен обзор интеллектуальных технологий, которые находят применение в социальном управлении, угрозы, связанные с практическим использованием интеллектуальных инструментов автоматизации и поддержки принятия решения, пути и меры снижения некоторых рисков, связанных с указанными угрозами.

Ключевые слова: технологическая безопасность, информационные риски, нейросетевые технологии, системы поддержки принятия решений, искусственный интеллект, цифровое управление персоналом.

Abstract: the article considers the intelligent automation of decision-making and management procedures that is being implemented in many areas of socio-economic practice, including financial and credit business processes, in trade and e-commerce (customer profiling, marketing micro-targeting), telecommunications, industry (technological control, robotics, neurocontrol, strategic planning and forecasting), intelligent automation also came to business management, to public administration. It is claimed that automation of personnel management is expanding (monitoring compliance with requirements, profiling and assessing KPIs, predicting conflicts and violations), unmanned vehicles and other neural network automation are used in medicine, the transport industry and agriculture; smart technologies come to education (in Moscow, a system of predictive analytics of the digital footprint of students is being tested to optimize and target educational services, help orientate in the future profession). The use of cognitive technologies in the creation of expert, advisory systems, decision support systems provides not only convenience and savings in time and effort, but gives rise to a variety of organizational, economic, ethical, social problems, giving rise to new risks. This study provides an overview of intelligent technologies that are used in social management, threats associated with the practical use of intelligent automation tools and decision support, ways and measures to reduce some of the risks associated with these threats.

Keywords: technological security, information risks, neural network technologies, DSS (decision support systems), artificial intelligence, people analytics, workforce analytics.

В последнее десятилетие успехи когнитивных и цифровых технологий усилили стремление компаний и государственных структур автоматизировать свои управленческие функции. Интеллектуальные технологии находят применение при решении самых разных многих задач управления: при принятии

управленческих решений разного рода, при прогнозировании и планировании, для осуществления контрольно-надзорных функций (в т.ч. для сбора, анализа данных об управляемых объектах), для автоматизации подготовки электронной документации, автоматизации управления деловыми, технологическими, логистическими и т. д. процессами. По мнению экспертов [10, с. 34–с. 34], до 90% всех задач по контролю, принятию решений, обеспечению коммуникации будет решаться с помощью инструментов ИИ уже в ближайшие 3 года. Интеллектуальные экспертные и советующие системы находят применение в работу органов муниципального управления (один из примеров: в московском ЖКХ реализуется пилотный проект прогнозирования аварийности и технического состояния жилых многоквартирных домов [13]), в работе министерств и ведомств. ИИ-инструменты поддержки принятия решения внедряются в медицину (для сбора и анализа данных, генерации контента, диагностики и терапии; так, в 2020 г. Департамент здравоохранения Москвы анонсировал эксперимент по внедрению интеллектуального анализа медицинских изображений в качестве средства поддержки принятия врачебных решений (см., например, [15])). В промышленности интеллектуальные системы используются для прогнозирования спроса, планировании запасов, управлении рисками, управление оборудованием и технологическими процессами (роботизация процессов, RPA), прогнозировании рыночных цен, контроле качества, оптимизации бизнес-модели. В прогнозе компании Meticulous Research [27] отмечает, что рынок промышленного ИИ составит 27 млрд. долларов в 2027 г., нефтегазовой отрасли – 3,81 млрд. долларов к 2025 г [25] и т. д.

Одна из сфер, где представляется перспективным внедрение интеллектуальной автоматизации – государственное управление. На протяжении последнего десятилетия в мире реализуется концепция «цифрового государства», в последние годы речь ведется о реализации «умного госуправления» (smart government). Инициаторами цифровой трансформации экономических, политических, социальных институтов стали IT-корпорации, транснациональные финансовые компании, затем были вовлечены административные элиты государств.

Содержание цифровой трансформации состоит в том, что реализуется государственно-частное партнерство, в госуправлении внедряются принципы проектного управления, другие элементы корпоративной управленческой культуры, выстраиваются корпоративные отношения между государством, обществом и IT-бизнесом, складываются своеобразные формы государственного «аутсорсинга» (когда часть контрольно-распределительных функций и полномочий государства делегируется частным IT-компаниям и финансовым корпорациям). Растет влияние технократии на содержание госпрограмм, а корпоративные цели и интересы «цифровиков» и финансистов включаются в стратегии государств в качестве основных направлений национального и экономического развития; создается благоприятная именно для IT-корпораций нормативно-правовая и институциональная среда; реализуется госфинансирование бизнес-проектов по внедрению цифровых платформ и экосистем. Растет и автоматизация государственного управления (при оказании госуслуг населению, коммуникации с ним, при выполнении фискальных и контрольно-надзорных функций, при сборе, интерпретации и обобщении сведений о социальных и экономических объектах). В настоящее время осуществление государством основных функций становится практически невозможным без использования информационно-коммуникационных инструментов.

Если взглянуть на историю реализации концепции цифрового государства (e-Government), концепция оформилась с 2000 года как корпоративная модель государственного управления, отличающейся от прежней модели госуправления большей «децентрализованностью», расширением участия бизнеса и ответственности, обновлением организационных структур. «Электронное правительство» трактовалось в начале 2000-е гг. как реализация комплекса бизнес-процессов, поддерживаемых информационными технологиями, которые позволяют информации беспрепятственно протекать между государственными учреждениями и программами и обладают интуитивно понятным интерфейсом для предоставления высококачественного обслуживания гражданам по всем государственным программам, во всех областях деятельности государства. Одним

из первых этапов стало внедрение систем безбумажного (электронного) документооборота и цифровых платформ дистанционного взаимодействия между административными органами, гражданами и экономическими субъектами. К сбору данных, предоставлению госуслуг были привлечены частные IT-компании, банки, бизнес-структуры. В 2009 США предложили странам инициативу «Open Government» («открытое правительство») как соединение информационно-коммуникационных технологий, принципов «прозрачности», «участия и сотрудничества» общественных институтов с политическим и экономическим управлением (Open Government: A progress Report to the American People, 2009). Был создан альянс США с рядом стран Европы и третьего мира OGP (Open Government Partnership). К 2011 году международные программы по расширению демократической и технологической составляющих электронного государства трансформировались в парадигму «smart government» («умное правительство», «умное управление»), – внедрение в государственное управление бизнес-процессов, реализуемых посредством цифровых технологий. Содержание концепции «цифрового управления» и «умного государства» раскрыто в работах таких авторов, как Kliksberg (2000), Willke (2007), Roberts, Grover (2012); Mellouli, Luna-Leyes, Zhang (2014), Awoleye, Ojuloge, Ilori (2014), Scholl, Scholl (2014), Rubel (2014), von Lucke (2016), Gil-Garcia, Zhang, Puron-Cid (2016), Bell (2017), Guenduez (2017), Schedler (2018)). Анализ двух десятков трактовок «e-Government», «Smart Government» за период 2000–2019 гг [1] не позволяет выявить четких политических или экономических характеристик этих терминов, что, полагаем, говорит об идеологической природе концепции «умного государства» («умного госуправления»), инструментарий которой используется для переопределения традиционной роли государства в социально-экономическом и политическом управлении, свидетельствуя о стремлении IT-корпораций и международных банковских структур перераспределить экономическую и политическую власть в свою пользу.

Как и другие страны, но с некоторым отставанием, Россия оказалась вовлечена в процессы цифровой реорганизации государственного управления, ре-

ализуется Национальная стратегия развития цифровой экономики в РФ. Принят ряд указов и федеральных законов в поддержку цифровой трансформации экономической, социальной сфер, государственного администрирования (в т. ч. ФЗ №8 от 9.02.2009 «Об обеспечении доступа к информации о деятельности государственных органов и органов местного самоуправления», ФЗ №210 от 27.07.2010 «Об организации предоставления государственных и муниципальных услуг», Указ Правительства РФ от 9.05.2017 №203 «О стратегии развития информационного общества в России на 2017–2030 гг.», Перечень поручений Президента по вопросам реализации нацпрограммы «Цифровая экономика РФ» (от 28.02.2019) и ПП РФ от 2.03.2019 №234 «О системе управления реализацией нацпрограммой «Цифровая экономика РФ», ФЗ №168 от 8.06.2020 «О едином федеральном информационном регистре, содержащем сведения о населении», ФЗ №258 от 31.07.2020 об экспериментальных правовых режимах в сфере цифровых инноваций в РФ и т. д.). Реализуются программы развития цифрового правительства («государство как сервис», «умное (цифровое) управление»), внедрение цифровых удостоверений личности. В Указе Президента РФ №474 от 21 июля 2020 г. «О национальных целях развития Российской Федерации на период до 2030 года» одна из целей – цифровая трансформация, заключающаяся в достижении цифровой зрелости ключевых отраслей экономики, здравоохранения, образования, государственного управления. При участии банков и IT-компаний, Министерства цифрового развития РФ разработана и утверждена национальная стратегия развития искусственного интеллекта и робототехники, предполагающая разработку и внедрение средств оперативной и прогнозной аналитики, рекомендательных и интеллектуальных систем поддержки принятия решений на основе технологий машинного обучения для анализа больших массивов данных, распознавания и синтеза речи, «умной» роботизации деловых процессов). Разработаны планы и дорожные карты по внедрению в работу министерств и ведомств РФ технологий интеллектуальной автоматизации.

Лидерами в технологической гонке по разработке технологий, инструментов ИИ и их внедрению являются США, Китай, Израиль. При этом, по оценке

экспертов [24, с. 12], в «копировании известных технологических решений и перенос на российскую почву» в России уже «нет конкурентных преимуществ для завоевания значимых сегментов мировой цифровой экономики... вместо этого Россия могла бы позиционировать себя, как «мировое конструкторское бюро» в области машинного интеллекта, сосредоточившись на решении фундаментальных проблем и взятии технологических барьеров». К слову, госфинансирование ИИ в России за 2017–2020 гг. составило около 8,7 млрд ₽ (в 350 раз меньше того же Китая) [5, с. 24]. В России с 2019 года действует «Национальная стратегия развития искусственного интеллекта до 2030 года». Среди задач развития искусственного интеллекта в России в Стратегии обозначены «повышение доступности и качества данных... для развития технологий ИИ», «создание комплексной системы регулирования общественных отношений, возникающих в связи с развитием и использованием технологий ИИ», поставлена и задача создания «универсального (сильного) ИИ», способного автономно решать различные задачи, автоматически обучаться, выполнять обработку информации на основе новых типов вычислительных систем, объяснять результаты обработки данных (интерпретируемый ИИ). ИИ должен стать частью процессов государственного и муниципального управления. В конце 2020 года утверждена программа по созданию в РФ стандартов технологии искусственного интеллекта на 2021–2024 гг., включая стандарты безопасности применения ИИ в транспорте, медицине, образовании, строительстве, стандарта «Умное производство» и др. Осенью 2020 года Указом Президента РФ утверждена «Национальная стратегия развития ИИ в РФ до 2024 года», где подчеркивается сквозной характер внедрения технологий ИИ, ставится задача «создания комплексной системы регулирования общественных отношений, возникающих в связи с развитием и использованием технологий искусственного интеллекта». Среди условий выполнения задач Стратегии – «формирование комплексной системы безопасности при создании, развитии, внедрении и использовании технологий искусственного интеллекта»; принципы, которые должны быть соблюдены при внедрении ИИ в практику управления – «обеспечение защиты прав и

свобод человека, в том числе права на труд», «предупреждение и минимизация рисков возникновения негативных последствий использования технологий ИИ», «объяснимость работы искусственного интеллекта», «недискриминационный доступ пользователей к информации о применяемых алгоритмах работы искусственного интеллекта», «преимущественное использование отечественных технологий и технологических решений на основе ИИ». Обозначенные в Стратегии области внедрения ИИ: «автоматизация рутинных производственных операций; использование автономного интеллектуального оборудования, интеллектуальных систем управления логистикой», «повышение лояльности потребителей», «оптимизация процессов подбора и обучения кадров», в сфере образования – «адаптация образовательного процесса к потребностям обучающихся и потребностям рынка труда, оптимизация профессиональной ориентации и раннее выявления детей с выдающимися способностями, автоматизация оценки качества знаний и анализа информации о результатах обучения», в медицине – «прогнозирование возникновения и развития заболеваний, подбор оптимальных дозировок лекарственных препаратов, сокращение угроз пандемий, автоматизация и точность хирургических вмешательств». Одна из задач Стратегии – «создание универсального (сильного) искусственного интеллекта, включая реализацию следующих приоритетов: а) алгоритмическая имитация биологических систем принятия решений, в том числе распределенных коллективных систем, таких как пчелиный рой или муравейник; б) автономное самообучение и развитие адаптивности алгоритмов к новым задачам; в) автономная декомпозиция сложных задач, поиск и синтез решений» [Указ Президента РФ от 10 октября 2019 г. №490 «О развитии искусственного интеллекта в Российской Федерации». – <http://publication.pravo.gov.ru/Document/View/0001201910110003>]. Также в 2020 г. в России утверждены Федеральный проект «Искусственный интеллект» и «Концепция развития регулирования отношений в сфере технологий ИИ и робототехники», принятие которых обусловлено (как сказано в документах) «серьезными вызовами перед правовой системой Российской Федерации, системой государственного управления и обществом в це-

лом», которые ставит технология ИИ, обусловленные «автономностью действий систем искусственного интеллекта в решении поставленных задач и их неспособностью непосредственно воспринимать этические и правовые нормы, учитывать их при осуществлении каких-либо действий» и отсутствием «законодательного регулирования, учитывающего специфику применения технологий ИИ и робототехники» [17]. В Концепции развития регулирования отношений в сфере технологий ИИ и робототехники» внимание уделено развитию исследовательских центров в области «сильного» искусственного интеллекта, доверенных систем, а также этических аспектов применения ИИ-технологий; при этом до 2030 в концепции предусмотрена разработка решений ИИ, аналогичных или «превосходящих человеческие показатели по широкому кругу задач. При этом IT-корпорациям и цифровым стартапам, готовым создавать и внедрять инструменты ИИ, предоставляется режим правового благоприятствования: к задачам регулирования в сфере ИИ в Стратегии отнесено (пункт 49): «а) обеспечение благоприятных правовых условий (в том числе посредством создания экспериментального правового режима) для доступа к данным, преимущественно обезличенным, включая данные, собираемые государственными органами и медицинскими организациями», «в) создание правовых условий и установление процедур упрощенного тестирования и внедрения технологических решений, разработанных на основе ИИ, а также делегирования информационным системам, функционирующим на основе искусственного интеллекта, возможности принятия отдельных решений (за исключением решений, которые могут ущемлять права и законные интересы граждан), в том числе при исполнении государственными органами государственных функций (за исключением функций, направленных на обеспечение безопасности населения и государства)», ж) разработка этических правил взаимодействия человека с искусственным интеллектом» [21].

В последние годы и месяцы российская законодательная система реформируется под задачи «цифровизации», причем цифровизация самого законотворческого процесса рассматривается как одно из институциональных условий

реализации цифровой экономики. В 2017 году Департамент проектной деятельности Правительства РФ предложил «Концепцию поэтапной цифровизации правовой системы», позже в Сколково в ходе конференции LegalTech профессионально обсуждались перспективы цифровизации права в России, Центр компетенции по нормативному регулированию цифровой экономики из Фонда «Сколково» стал главным разработчиком документов нормативного регулирования цифровой экономики. Нормативное регулирование использования ИИ в России в настоящее время определяется содержанием Национальной стратегии, национальной программы «Цифровая экономика Российской Федерации», Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы, «Концепцией регулирования ИИ и робототехники до 2024 года» и др. В России ИТ-компании, финансисты, АНО «Цифровая экономика» привлечены к правовому регулированию использования ИИ; в 2019 г. фондом «Сколково», ПАО «Сбербанк России», ПАО «МТС», Аналитическим центром при правительстве РФ, Национальной ассоциацией участников рынка робототехники, Центром стратегических разработок, компаниями Яндекс и Mail.ru Group инициировали проект законодательной концепции регулирования ИИ (в т.ч. вопросы обработки данных, распределения ответственности, прав на результаты машинного творчества, вопросы применения искусственного интеллекта и робототехники в транспорте, медицине, образовании, вопросы безопасности и персональных данных пользователей, взаимодействующих с искусственным интеллектом).

К сожалению, в документах развития ИИ в РФ не обозначены механизмы контроля за развитием технологий ИИ и безрисковым внедрением в практику управления ИИ-инструментов, предполагающих автономию от человека. Еще одно положение Стратегии связано, на наш взгляд, с рисками для объектов критической информационной инфраструктуры и граждан, данные о которых должны будут собираться и предоставляться для обучения разработчикам коммерческих инструментов ИИ: «Основными направлениями повышения доступности и качества данных, необходимых для развития технологий искусственно-

го интеллекта, является создание (модернизации) общедоступных платформ для хранения наборов данных (в том числе звуковых, речевых, медицинских, метеорологических, промышленных данных и данных систем видеонаблюдения) для обеспечения потребностей организаций – разработчиков в области искусственного интеллекта»; «к 2030 году объем опубликованных на общедоступных платформах наборов данных должен быть достаточным для решения всех актуальных задач в области ИИ, в том числе за счет публикации звуковых, речевых, медицинских, метеорологических, промышленных данных и данных систем видеонаблюдения» [21] т. о., создание и обучение ИИ предусмотрено с использованием реальных данных граждан, размещенных на «общедоступных платформах», – а не синтетических данных, что было бы безопаснее. К сожалению, в документах стратегического уровня практически не упоминается об этической стороне предоставления широких полномочий разработчикам ИИ в ущерб интересам граждан, поставленных в роль цифровых «морских свинок» по отношению к разработчикам моделей ИИ. Между тем, цели разработчиков ИИ, получающих доступ к данным граждан через «общедоступные платформы», могут быть различны, включая узкокорпоративные коммерческие, политические, военные, недружественные и даже криминальные, но проконтролировать на практике законность и этичность дальнейшего использования, обученного ИИ и данных сложно. Между тем такие инструменты ИИ, обученные на массиве медицинских, поведенческих, социально-демографических, финансовых и др. данных с «общедоступных платформ», могут использоваться для коммерческого профилирования и розыска людей с учетом их биомедицинских, психологических, политических, профессиональных особенностей, использовать эти дата-сети для силовых и военных задач, для социальной инженерии (социального скоринга, управления поведением), чему есть немало примеров в практике IT-компаний. Процесс создания моделей ИИ облегчается и благодаря тому, что корпоративные разработчики инструментов ИИ (Microsoft, Google, Facebook, Nervana, Baidu и т. д.) и университеты публикуют в открытом доступе библиотеки исходного кода и программные системы для создания и обуче-

ния нейросетей (примеры общедоступных ИИ-инструментов: Caffe, Theano, TensorFlow, Torch, CNTK, Deepmind Google, scikit-learn, DMTK, Deeplearnig4j, H2O, Mahout, MLLib, PaddlePaddle, MXNet, NuPic, и др.) [см. например, 19]. Все это облегчает доступ к технологии разработчикам с «пограничными» и криминальными целями, размывает ответственность за последствия использования технологии в отношении национальной, экономической, личной безопасности. Торговля поведенческими продуктами – нейросетями, обученными на данных реальных людей, способными прогнозировать поведение людей и управлять им, – приносит брокерам данных (Acxiom, Oracle, Experian, Epsilon, банки и т. д.) и IT-компаниями (Facebook, Google, Microsoft и т. д.) десятки миллиардов долларов ежегодно. Цифровые аналитические ИИ-инструменты и поведенческие технологии в сочетании с массивами открытых данных, позволяет решать и политические задачи (подобно модели OKEAN для анализа личных данных из соцсетей от компании Cambridge Analytica, позволивших заинтересованным лицам манипулировать через соцсети поведением избирателей во время выборов в США, Великобритании, Кении) [4]. В стандартах безопасности и нормативных документах говорится, что снизить риски незаконного использования собираемых сведений о гражданах могли бы регулирование доступа к хранилищам таких данных и операции обезличивания (деидентификации) данных. Однако, как показывает практика, данные имеют свойство «утекать» на черный рынок, оказываются у брокеров данных, которые находят этой «новой нейти» разнообразное коммерческое применение. ГОСТ Р 55036–2012 ISO/TS 25237:2008 «Информатизация здоровья. Псевдонимизация» констатирует, что «никакая процедура де-идентификации не может... гарантировать сокрытие идентичности». По свидетельству специалистов по информационной безопасности, операции обезличивания и шифрования персональных данных перед размещением на «общедоступных платформах» во многих случаях все равно не становятся препятствием для последующего восстановления идентичности субъекта данных, т.е. для идентификации личности по его «обезличенным» персональным данным из разных источников, и это доступно генеративным

нейросетям [30]. Хотя в России идет работа над законопроектом, запрещающим деобезличивание данных (кроме случаев защиты жизни или здоровья человека), фактически технология деобезличивания существует, поэтому конфиденциальные данные о гражданах на общедоступной платформе могут быть стать источником махинаций и угроз национальной безопасности. Наилучшим с точки зрения безопасности было бы просто не предоставлять широких свобод разработчикам ИИ и коммерческим интересантам в отношении данных россиян, в ущерб праву на приватность и контроль за использованием личных сведений; следовало бы хранить данные децентрализованно, а не связанными воедино через сводный реестр (ЕФИР, ЕСИА и т. п.), не отчуждать без согласия граждан сведения о них, а сделать хранение и использование своих данных более подконтрольным для граждан (технологии для этого существуют).

Искусственный интеллект, согласно определению ГОСТ 33707–2016 «Информационные технологии. Словарь», это «способность функционального блока выполнять функции, обычно ассоциирующиеся с интеллектом человека, такие, как, например, рассуждения и обучение». Привычное определение искусственного интеллекта трактует его как деятельность, направленную на создание «думающих» искусственных систем, или систем, способных функционировать адекватно среде, просчитывая стратегию. Сегодня технологии искусственного интеллекта представляют собой достаточно широкий круг технических и программно-технических систем, включая специализированные суперкомпьютеры, робототехнику, когнитивные компьютеры, нейроморфные чипы (Loichi и т. п.), байесовы модели, нейронные сети (сверточные, капсульные, рекуррентные, генеративно-состязательные) в сочетании с методами глубокого машинного обучением на массивах данных и знаний). Сферы и отрасли, где находит применение искусственный интеллект – промышленность (автоматизация технологических процессов), финансы (прогнозирование и анализ рисков, банкинг, биржевые операции), масс-медиа, медицина, решение разнообразных задач бизнеса, таргетирование рекламы, рекомендательные системы, машинное зрение, распознавание речи, прогнозная аналитика и принятие решений, безопасность. ИИ-

технологии позволяют автоматизировать или имитировать отдельные когнитивные и физические функции, решая формально поставленную узкую задачу (например, распознавание или генерация контента (так, модели BERT, GPT-2, GPT-3, архитектура Transformer способны распознавать семантику и смысл, генерировать простые тексты на заданные темы), выявление закономерности в больших или зашумленных массивах данных, поиск оптимального в некотором контексте решение, прогнозирование по определенной методике, кластеризация и профилирование объектов контроля и управления и т. д.). Основные тренды развития интеллектуальных технологий поддержки принятия решений: автоматизированная разметка данных, потоковая обработка данных (real time data streams), автоматизации и полуавтоматизации машинного обучения (Automating machine learning, AutoML). Инструменты для создания когнитивных систем разнообразны (это интеллектуальные алгоритмы поддержки принятия решений для FuzzyCalc, CubiCalc, BrainMaker, IDIS, BERT, ISIS и экспертных систем EXSYS, GURU, KL, KRYPTON и проч., основанные на правилах, фреймах, иных формализованных знаниях; это системы имитационного моделирования (NanoCAD и др.); библиотеки и системы машинного обучения (Caffe, Theano, TensorFlow, Torch, CNTK), фреймворки для оптимизации «черного ящика», black-box optimization (BlaBoO, Kurobako, COmparing Continuous Optimisers) и проч., языки и системы программирования (в т.ч. вероятностного программирования и моделирования) и т. д.)

«Разумность» инструмента ИИ выражается в математической логике или статистической близости результатов численной оптимизации к образцу. Примечателен «парадокс», сформулированный робототехником и футурологом Х. Моравеком: высококогнитивные процессы требуют значительно меньших вычислений и усилий по их моделированию, чем имитация низкоуровневых сенсорных, связанных с ощущениями физического мира и способностью в нем ориентироваться. Но парадокс отражает лишь специфику подхода IT-специалистов к моделированию поведенческой и интеллектуальной «функции» человека (интеллект в моделях ИИ, как правило, сводится к формально-

логической, вычислимой рациональности). Насколько хорошо такая модель разума человека способна решать вместо человека задачи реального мира – вопрос открытый; как хорошо известно, обученные определенной функции программы, нейросети или роботы, теряют способность выполнять ее правильно, если обстановка изменилась. Любая система ИИ является упрощенной математической моделью биологических «систем принятия решений», причем модель вбирает в себя все ограничения, присущие целям, методу, условиям, инструментам моделирования, обучающему набору данных. Например, системы поддержки принятия решения и советующие системы на основе продукционных правил, других моделей формализации экспертных знаний просто использует логику рассуждения (или готовые решения) в соответствии с системой критериев и базой знаний (формализованного опыта специалиста-человека); модели глубокого обучения вырабатывают собственное «правило», обусловленное конкретным массивом обучающих данных и архитектурой нейросети. Не обладая собственной витальностью и долгим эволюционным опытом живых систем, ИИ-инструмент решает задачу формально и в узком контексте, по правилам, не принимая в расчет «неявных параметров» среды, которые разработчик не заложил в модель; для человеческого же мозга неявные параметры и сигналы являются неосознанной и неформализованной частью решаемых задач, а приоритеты и критерии оптимальности меняются с изменением окружающей среды. Искусственному интеллекту в силу его глубокого отличия от человека, не свойственны инстинктивные или этические основы поведения. Поэтому управленческое решение задачи социального управления, найденное «интеллектуальным инструментом», не обязательно будет верным с точки зрения человека, решающего ту же задачу с учетом широкого социального контекста, опыта, последствий, других неявных обстоятельств. Основное достоинство инструментов принятия решений – быстрота, – достигается упрощением модели объекта анализа или управления, за счет формализма и отсекающей значительной части контекстной информации об объекте, что делает малоэффективным использование ИИ в управлении «на длинной дистанции» и в управлении сложными и само-

развивающимися объектами. По этим причинам, полагаем, современный «слабый ИИ» – однозадачный искусственный интеллект, – неэффективен в качестве средства автономной автоматизации управления сложными социальными и биологическими системами; и ограничено пригоден для выполнения отдельных управленческих задач в отношении объектов, сравнимых с ИИ-моделью по сложности (т.е. для простых искусственных систем, технических, виртуальных объектов, не способных к развитию или предполагающих контролируемую динамику в пределах заданных значений параметров).

Ряд авторов (В. Виндж, Р. Курцвейл, И. Маск и др.) прогнозирует к 2045 году (или раньше) появление «сильного (универсального) ИИ» (Artificial General Intelligence), обладающего способностью, подобно человеку, самостоятельно ставить и решать комплексные интеллектуальные задачи в разных сферах деятельности, возможно, обладающий самосознанием и собственным целеполаганием. В последние годы около 3 десятков государств (США, Германия, Англия, Австралия, Италия, Дания, Финляндия, Франция, Швеция, Канада, Индия, Китай, Япония, Сингапур, Мексика, Тунис, ОАЭ, Кения и др.) сформулировали национальные стратегии развития искусственного интеллекта; государства соревнуются за первенство в этой отрасли и ставят задачу создания универсального («сильного»). Р.Курцвейл назвал этап появления универсального AGI «технологической сингулярностью», полагая, что в этом случае человечество уступит роль доминирующего «вида» на земле «сильному ИИ». По мнению прогнозиста С. Переслегина [16] сильный ИИ будет способен манипулировать каждым жителем планеты индивидуально, меняя его жизнь под свои задачи. Многие исследователи ИИ полагают, что следует уже сейчас на государственном уровне озаботиться созданием условий, которые бы обеспечили безопасное для социума развитие интеллектуальных технологий. Илон Маск (чей стартап Neuralink известен созданием чипа нейроинтерфейса и робота-«нейрохирурга») выразил опасения, что люди могут оказаться в роли биологической платформы, субстрата для становления цифрового супер-интеллекта; по мнению Маска, искусственный интеллект – потенциально более опасная техно-

логия, чем ядерное оружие. Специалист в области интеллектуального анализа данных и технического зрения Ю. Визильтер [2] отнес ИИ и робототехнику в «токсичным технологиям», столь же опасным по своему воздействию на среду и социум, как химическое производство, ядерная энергетика, телевидение, интернет и множество других современных технологий, с которыми приходится мириться, пока польза от их применения превышает наносимый ими вред. В п. 9 «Национальной стратегии развития ИИ в РФ до 2030 года» также отмечается, что «создание универсального (сильного) искусственного интеллекта... может привести не только к позитивным изменениям в ключевых сферах жизнедеятельности, но и к негативным последствиям, вызванным социальными и технологическими изменениями» [Указ Президента Российской Федерации от 10.10.2019 №490 «О развитии искусственного интеллекта в Российской Федерации» <http://ips.pravo.gov.ru:8080/default.aspx?pn=0001201910110003>].

ИИ все шире используется как инструмент контроля над людьми (в управлении персоналом, контроле за гражданами, преступниками, в прогнозировании поведения); ИИ применяются для автоматизации принятия решений о финансовой и правовой участи граждан и юридических лиц. В управлении инфраструктурой, данными, социальными объектами уже используются сотни нейросетей. В цифровой медицине ИИ-инструменты применяются для диагностики и выработки медицинских рекомендаций (Watson AI, BotkinAI, и т. п.). В масс-медиа, рекламе ИИ-инструменты помогают заинтересованным группам формировать коллективное сознание, направлять внимание и поведение, генерируя и фильтруя контент цифровых и традиционных СМИ, ранжируя политические и научные публикации в новостных и поисковых выдачах глобальных сетей, помещая людей в «информационные пузыри», формируя убеждения, социальный опыт, привычки и ценности благодаря «цифровым платформам», поведенческим и интеллектуальным технологиям. Интеллектуальные системы уже несколько лет участвуют в принятии судебных решений; полиция использует системы автоматического распознавания лиц прохожих для опознания правонарушителей и принятия решений об аресте подозреваемых; алгоритмы

оценки рисков выносят рекомендации о желательной мере пресечения для лиц, находящихся под следствием. Например, американская полиция и банки применяют программно-аппаратный комплекс Palantir (способную вычислять людей, которые могут совершить насильственное преступление или стать его жертвой) [9], существует британская ИИ платформа «DSAAS (Police)», во Франции, Японии, Голландии в силовых ведомствах для расследования некоторых преступлений применяются ИИ-платформы прогнозной аналитики, в ЕС с 2013 г. сотрудники Европола используют систему разведки и раннего обнаружения преступлений международных ОПГ ePOLICE.

В России для развития цифрового госуправления региональным органам управления в РФ поставлена цель повышения эффективности деятельности органов власти за счет внедрения платформенных решений на базе технологий искусственного интеллекта», достижение которой связывается с решением таких задач, как кардинальный реинжиниринг госуслуг и сервисов, дистанционный мониторинг объектов контроля, роботизация рутинных операций, автоматический сбор данных с помощью интернета вещей. Проекты по использованию ИИ в федеральных органах власти РФ осенью 2020 года представили Министерство цифрового развития, МВД, Министерство чрезвычайных ситуаций, Федеральное агентство по делам молодежи, Федеральный фонд обязательного медицинского страхования [18], Министерство здравоохранения, Министерство транспорта, Министерство сельского хозяйства [7]. Федеральная налоговая служба, мировой лидер цифровизации налоговой функции, использует нейросетевого голосового ассистента по налоговому администрированию в личном кабинете физлица, ИИ используется при работе с электронными документами, планировании) [22]. Согласно дорожной карте «Нейротехнологии и искусственный интеллект» (разработанной Сбербанком при участии компаний, вошедших в «Альянс по развитию ИИ») [11]. Министерство цифрового развития РФ планирует до 2024 года внедрить решения на основе ИИ в МВД (опознание, анализ взаимосвязей между событиями, людьми), в МЧС (поиск пожаров, поиск инфраструктуры и разрушений по снимкам), в РосРеестре (поиск объектов

недвижимости). В Минздраве к 2024 г. половина медучреждений обязана внедрить ИИ-сервисы (для диагностики, интеллектуальной обработки данных с носимых устройств; роботизация рабочих процессов, в т.ч. генерации медицинских документов путем распознавания речи, планируется ввод ИИ-платформы экспертизы медицинской продукции, платформы с технологиями ИИ для оказания госуслуг, доступа к единой базе медицинских данных [8]. Внедрение ИИ-инструменты поддержки принятия решения ИИ запланировано в Минпромторге, в Минюсте (правового ассистента, ИИ-систему экспертизы нормативно-правовых актов, систему мониторинга правоприменения). Роснедра планируют применять систему поддержки принятия решений на основе ИИ и нейросетевого анализа, выполняющую «сбор, обработку, предоставление доступа к машиночитаемым дата-сетам геологической информации, необходимым для принятия управленческих решений», ИИ-помощника недропользователя. Росморречфлот будет использовать ИИ-систему сбора данных о движении судов с учетом погодных и навигационных условий, на данных будет самообучаться интеллектуальная рекомендательная система судоходства. В Роспотребнадзоре для автоматизации обработки обращений есть онлайн чат-бот с системой распознавания речи, запланирована разработка платформы для эпидемиологического моделирования и прогнозирования. По словам замминистра обороны РФ Р. Цаликова в 2018 г. на конференции «Искусственный интеллект: проблемы и пути решения», «отдельные элементы искусственного интеллекта или системы интеллектуального управления в Вооруженных силах активно применяются, например, в беспилотных системах и робототехнике» [3], планируется использовать ИИ в БПЛА, для боевых задач без непосредственного приказа человека в момент полёта, автономного поиск отдельных типов целей [12] и т. д. Таким образом, практика и планы внедрения ИИ охватывают практически все структуры российского госуправления и многие аспекты взаимодействия государства и общества. Согласно последней версии федерального проекта «Искусственный интеллект» к 2024 г., около 5,5 сотен малых предприятий и более 60 коллекти-

вов разработчиков получают господдержку в создании ИИ-инструментов и реализации ИИ-инноваций.

Среди преимуществ когнитивной автоматизации управления социально-экономической инфраструктурой – снижение доли рутинной деятельности, прозрачность и подконтрольность управляемых процессов. Использование ИИ экономит ресурсы за счет сокращения участников согласования при принятии решений, за счет сокращения персонала и упрощения объекта управления, – управлять людьми сложнее, чем автоматикой. – Однако сама ИТ-технология требует больших затрат – финансовых, организационных, человеческих и интеллектуальных, материальных (элементная база, которая постоянно совершенствуется, требует много редкоземельных металлов, энергоресурсов для своего развития и функционирования), создавая отток всех этих ресурсов из «нецифровых» отраслей и сфер. Еще одно достоинство ИИ в том, что при решении ряда управленческих задач ИИ действует менее предвзято и более формально, чем человек, принимающий решение, что снижает коррупционную составляющую управления. Среди негативных сторон использования «интеллектуальной» автоматизации управленческих и т. д. задач – риск предвзятости ИИ (обусловленной особенностями модели ИИ), неподконтрольность «логики» принятия решений инструментом ИИ, неточность самих решений; размывание ответственности за последствия решения или действия «цифрового субъекта»; риски несанкционированного доступа к конфиденциальным данным, которые используются при создании и деятельности модели ИИ; формализм решений ИИ и цифровое выхолащивание социальных отношений при использовании роботизированной автоматизации процессов. Искусственные системы управления не понимают и не учитывают социобиологических свойств человека, управляют людьми как техническими объектами, зачастую апеллировать из-за неприемлемого стиля управления под управлением интеллектуальных систем, нарушения прав или несправедливости решения ИИ просто некуда, поскольку управляет не человек, а «интеллектуальный» автомат. Пример российский: в последние годы в России частные компании, используя системы и камеры автоматической ви-

деофиксации, связанные с нейросетевой системой распознавания, получают немалую прибыль, штрафуя нарушителей ПДД; при этом велико количество ошибок и спорных ситуаций (например, настройка камер так, чтобы фиксировать превышение скорости автомобилей, когда она не превышена). Самая распространенная ошибка таких систем распознавания – ошибочное распознавание госномера, когда автомобиль-нарушитель один, а номера, зафиксированные и распознанные «умной камерой», принадлежат другой машине, или штраф за превышение скорости приходит автомобилю, который ехал на эвакуаторе. Хотя водитель может опротестовать штраф в ГИБДД, доказать невиновность непросто (требуется либо предоставить запись с видеорегистратора, либо в документе фотофиксации должно очевидное противоречие). Т. о., люди оказываются в более уязвимом положении по сравнению с автоматизированными системами.

Ряд интеллектуальных инструментов поддержки принятия решений реализован на основе модели «черного ящика» (т.е. не объясняют пользователю логику принятия решения и никак иначе не визуализируют процедуру получения результата, что вынуждает ЛПР брать на веру достоверность вывода); при этом зачастую выносимые решения относятся к «чувствительным» для людей финансовым, юридическим вопросам, касаются свобод или карьерных возможностей, вопросов безопасности. Компании по предиктивной аналитике разрабатывают ИИ-инструменты, позволяющие на основе статистических оценок вероятности событий прогнозировать социальные процессы, управление которыми на основе таких инструментов удобно, но ставит этические вопросы и требует более веских обоснований, чем статистическое усреднение. Например, при реализации концепции умного города на практике подобные системы могли бы помочь в решении ряда задач (оценка состояния инфраструктуры, преступности, прогнозирование развития эпидемиологической обстановки и вообще прогноз здоровья и проч.). Так в Рио-де-Жанейро есть опыт внедрения системы CrimeRadar для прогнозирования криминогенной обстановки [28]. В более чем в 70 странах городские службы «умных городов» и полицейские на практике используют те или иные виды предиктивной аналитики [14], ИИ-инструменты

предиктивной аналитики (например, с 2014 г. – COMPAS) применяются в полиции США (штаты Калифорния, Вашингтон, Южная Каролина, Алабама, Аризона, Теннесси, Нью-Йорк, Иллинойс), в Германии (инструмент Precobs [20]), в Евросоюзе (ePOOLICE); в России также рассматривается возможность применения систем предиктивной аналитики. Применение подобных инструментов ставит технологические и этические вопросы, формальные статистические инструменты явно недостаточны для обоснования судебно-правовых решений, даже если они удобны и законодательно оправданы. Несколько лет в американских судах для оценки поведения обвиняемых в США использовалась советуемая нейросетевая система предиктивной аналитики Compas, вычислявшая риск совершения повторного преступления, влияя на решение о сроке заключения виновному. Алгоритм компании Northpointe предсказывал вероятность повторного преступления подсудимым в суде в округе Броуард (Флорида), помогая судье определить, наказать преступника; анализ результатов работы программы, но один случай в 2016 г. расследованный журналистами ProPublica обнаружил расовую предвзятость инструмента (его ошибочное предсказание рецидива для чернокожих было вдвое выше, чем для белых) [26]. Еще пример предвзятости и непрозрачности нейросетевых моделей при принятии решений: в 2014 году компания Amazon разработала систему искусственного интеллекта для найма персонала, но выяснилось, что ИИ годами делала выбор предвзято, отклоняя кандидатуры женщин и выбирая лишь мужчин, поскольку нейросеть была обучена на выборке, состоящей в основном из резюме мужчин [23].

Несмотря на специфику моделей ИИ, за ними зачастую признается презумпция невиновности, а объект, в отношении которого допущена ошибка, оказывается в роли виновного «по умолчанию». Разработчики ИИ-моделей для идентификации говорят о 2 видах ошибок, присущих любым подобным моделям, – ошибки 1-го рода (система не смогла идентифицировать известный ей объект) и 2-го рода (система ошибочно опознала непричастного, например, приняла чужого за своего). Если модель используется для вынесения решения о наказании, система не должна доказывать свою правоту, но при ее ошибках 2-

го рода потерпевший вынужден опровергать обвинение ИИ-системы, доказывая обратное. Автоматическая идентификация граждан с помощью нейросетей является популярным средством автоматизации правоприменительной практики, используются для распознавания правонарушений, поиска преступников, управления персоналом, – в подобных обстоятельствах ИИ-модели, действующие по принципу «черного ящика», чем создают риски нарушения прав граждан. Вот несколько примеров ошибочных решений с использованием нейросетевых технологий. До 2020 года полиция США использовала систему распознавания лиц Amazon Rekognition для поиска преступников; в 2019 году эта система распознала по фотографиям 28 игроков НФЛ как преступников, – к счастью, лишь на этапе тестирования. Другой пример: Москве в октябре 2020 г. автоматическая система распознавания лиц сети Ашан в магазине «Ашан Сокольники» «опознала» преступника в обычном покупателе на основании его сходства его облика с реальным вором («человек среднего телосложения в черной кепке, черной кофте, синих джинсах, с обычной белой медицинской маской на лице, из-под которой была видна борода»), и по указанию нейросети охранники магазина задержали и передали покупателя в полицию, не вникая в возражения задержанного; наняв адвоката, через несколько часов, потерпевший смог выйти из заключения, поскольку адвокат доказал полицейским необоснованность задержания лишь на основании обнаружения нейросетью сходства. Возмещения ущерба за ошибку со стороны службы безопасности гражданин не добился, – решение принял ИИ, а люди (охранники и полицейские) это решение выполнили автоматически [6]. Таким образом, наблюдается размывание ответственности за решение, если в цепочке исполнителей решения оказываются и люди, и интеллектуальные технологии. Подобные обстоятельства меняют социальные «правила игры» и некоторые принципы социального управления (например, делегирование и распределение ответственности за управленческое решение). Существует дефект технологии генеративных нейросетевых моделей, используемых в системах распознавания, каждый слой таких сетей, применяемых для распознавания, классификации и т. п., анализирует в предъявленном изображении отдель-

ные его элементы, а не образ в целом (в отличие от человека); если вместо изображения известного сети объекта предъявить беспорядочную мозаику частей этого изображения, нейросеть все равно «увидит» объект, а не предъявленную хаотическую картину, нейросети не «чувствуют» более широкий контекст, в котором принимается решение. В Китае система автоматического распознавания правонарушителей в 2018 г. она прислала штраф предпринимательнице Дун Минчжу потому, что фото оказалось на борту автобуса, проезжавшего пешеходный переход (на фото, сделанном системой фотофиксации нарушения, эта ошибка очевидна для любого человека, – для ИИ-системы, принявшей решение о штрафе, никакой ошибки нет).

С автоматизацией процессов управления связан риск технический сбоев, как намеренных, так и случайных (обусловленных методологиями разработки программного обеспечения – применением при разработке продуктов готовых фреймворков, библиотек, платформ разработки, генерирующих часть программного кода непрозрачно для разработчика и оставляющих лазейку (закладки и backdoor в готовом программной продукте) для потенциальных злоумышленников, спецслужб, экономического шпионажа. Инструменты разработки ИИ находятся в открытом доступе, и желающие могут создать свой нейросетевой инструмент любого назначения; разработчики бесконтрольны со стороны общества или регуляторов в том, какие цели они перед собой при этом ставят и какие критерии будет учитывать инструмент ИИ. Источником рисков работы ИИ-инструментов могут стать недальновидность, социальная некомпетентность, предвзятость или безответственность разработчиков, заказчиков инструмента ИИ, злонамеренные действия (например, отравление обучающих или тестовых данных для нейросетевых инструментов). Примеры отравления обучающих данных: в ходе тестирования нейросетей в социальных средах пользователи обучили ИИ-ботов (Watson, Alisa) сквернословию. Другой пример: часто разработчики нейросетей для распознавания изображений обучают их на изображениях людей в соцсетях; в дальнейшем обученная сеть может опознать реального человека по фотографии, которая была в обучающем массиве (и как было

показано выше, может сделать ошибку 2-го рода при распознании), а применять подобные сети могут, например, для поиска преступников среди прохожих и т. д. Некий стартап создал сетевой сервис, использующий «отравление» данных как средство безопасности, предлагая пользователям соцсетей перед размещением своих изображений в открытом доступе вносить искажения с помощью сервиса; человек искажений не замечает, но нейросеть, обученная на «отравленных» фото из соцсетей, не сможет идентифицировать человека по такому его фото. Поскольку способность решать некую типовую задачу нейросетью фиксируется в ее архитектуре, то даже небольшие искажения в исходных данных или условиях задачи резко затрудняют ей поиск правильного решения. Например, ИИ Google Health, обученный для диагностики глазных заболеваний, в лаборатории разработчика ставил верные диагнозы в 90% тестового набора снимков, но в условиях реальных клиник Таиланда не мог распознать более 20% снимков пациентов [29].

Риски применения систем принятия решений критичны в военной сфере при решении задач с применением военных беспилотников, интеллектуальных систем принятия решения. Существует три уровня автономии для систем вооружений: оружие первого уровня автономии широко используется сегодня, требует человеческой команды для выбора роботом цели и развертывания силы; оружие второго уровня автономии само выбирает цели и развертывает силы, но человек может проконтролировать и отменить решения робота; оружие третьего уровня автономии – полностью автономное оружие, которое никак не зависит от человеческого контроля, выбирает цель и начинает атаку без участия человека. Считается, что применение беспилотной военной техники уменьшает риск несанкционированного использования оружия (удары борта БПЛА можно отследить), их обслуживание требует меньше ресурсов, чем использование традиционной военной техники. По мнению специалистов и военных, причастных к разработке военных системы на основе искусственного интеллекта, исключение людей из процесса принятия решений в военной обстановке позволит наносить меньше случайного, побочного ущерба, при условии, что такое оружие не выйдет

из-под контроля или не будет совершать ошибок, однако на практике военный ИИ ошибается. «Как свидетельствуют результаты исследования авиаударов беспилотников НАТО, в период с января 2009 г. по ноябрь 2013 г. было совершено более 120 авиааналетов. Их них 42 стали ошибочными, то есть привели к гибели мирных жителей. Всего в результате авиаударов погиб 1391 человек и только 905 из них оказались боевиками» [19]. Кто должен нести ответственность за ошибочные решения, принятые автономными искусственными агентами, определить бывает сложно, что дает большой простор для злоупотреблений и серьезных преступлений. Поэтому контроль над системами вооружений на основе компьютерных программ требует все же совместной работы человека и машины, чтобы были оптимизированы сильные стороны каждого из них, иначе сбои в работе военных беспилотников способны нанести большой урон.

Обществу приходится адаптироваться к тому, что роботы и технологии системы искусственного интеллекта становятся рабочим классом, в т. ч. частью системы управления и принятия критически значимы решений на всех уровнях. Но вопросы юридической субъектности ИИ (права и ответственность искусственных интеллектуальных субъектов или их владельцев) не решены, общие стандарты в отношении юридического статуса ИИ в России не сложились. Между тем интеллектуальная автоматизация уязвима к кибератакам. Автономный транспорт может не распознать как препятствие человека на дороге, одетого в камуфляжную форму, или распознать испорченные дорожные знаки (такие прецеденты возникают в процессе испытаний беспилотных автомобилей), следствием такого дефекта технологии или «отравленных» обучающих данных станет в будущем несчастный случай; кто окажется виноват – владелец управляемого объекта (автомобиля), или разработчик автопилота, или создатель интеллектуальной модели, или компания, выпустившая автомобиль-беспилотник, или потерпевший, или пассажир беспилотного автомобиля? Ответственность словно бы расплывается при использовании средств автоматизированного принятия решения, внося хаос в традиционную систему норм и практики. Но кто-то должен нести ответственность перед потерпевшими и обществом, чтобы со-

хранялась социальная стабильность и социальная управляемость. Наделение юридической ответственностью автономных и полуавтономных ИИ-систем принятия решения предполагает и юридическое признание определенных прав за ними. Но может ли нести ответственность технический объект, не обладающий самосознанием, и что считать самосознанием? Подобные вопросы остаются открытым и для своего разрешения требуют привлечения специалистов гуманитарных наук (философов, психологов, юристов), а не только технических.

Интеллектуальные технологии не только упрощают контроль управление, но меняют рынок труда; среди негативных социальных последствий интеллектуальной автоматизации – исчезновение целых профессий, рост технологической безработицы среди специалистов, чьи функции переданы ИИ и робототехнике. Следствия этого процесса – потеря источников доходов больших социальных групп, падение уровня их жизни, криминализация, социальная дестабилизация и деморализация, обусловленные лишением людей возможности вести трудовую деятельность. Для всех категорий наемной рабочей силы, кроме прямых выгодополучателей автоматизации, происходит разрушение «нецифровых» социальных практик. Будет расти и социальное неравенство из-за «технологических барьеров» (в силу территориальной неоднородности распределения технологической инфраструктуры, неспособности людей платить за цифровые средства доступа к благам цифровой экономики из-за отсутствия источников дохода, из-за физиолого-возрастных и др. ограничений доступности цифровых сервисов, через которые намерены в новых условиях работать с населением и «цифровое государство», и «цифровой бизнес».

Каковы могут быть пути снижения рисков, связанных с небезопасным использованием технологий ИИ? В Европе с 2018 года действует Европейская этическая хартия о применении искусственного интеллекта в судебных системах; в 2020 году Еврокомиссией была принята «Белая книга об искусственном интеллекте. Европейский подход к совершенству и доверию», но зарубежные и отечественные специалисты по безопасности часто отмечают, что нормативная деятельность государств и общественности отстает от развития технологий ИИ

и внедрения их в социальную практику. России еще предстоит разработать госстандарты в сфере разработки безопасного ИИ и робототехники.

Проблема массового вытеснения с рынка труда целых категорий специалистов-людей, чьи функции передаются средствам интеллектуальной (и не только) автоматизации будет с каждым годом обостряться. Исследование путей решения этой серьезной и острой социальной проблемы, конечно, выходит за рамки одной статьи; упомянем лишь одну концепцию. В качестве альтернативы трудовым доходам, которые станут многим недоступны по мере роста технологической безработицы, ряд экономистов обсуждает концепцию безусловного базового дохода. Она предполагает, что безработным и нуждающимся, которые не могут трудоустроиться в новых условиях, государство предоставляет некий минимум благ для поддержания существования, за счет налогов, которую платят, например, владельцы бизнесов, использующих вместо людей средства автоматизации, робототехнику; граждане, получающие безусловный базовый доход, могут быть привлечены к выполнению социально полезных и т. п. работ. Но в долгосрочной перспективе этот путь не решит ни проблем технологической безработицы, ни других социальных проблем, порожденных цифровизацией; экономика невостребованных населением товаров будет пребывать в кризисе, не находя спроса на продукцию, не развиваясь, экологические проблемы будут нарастать, поскольку технология производства и утилизации ИТ-продукции затратная и «грязная», а общество «лишних людей» рискует трансформироваться в рабовладельческую формацию или еще более деструктивные формы. Поэтому способ решения проблемы технологической безработицы должен быть другим (например, какое-либо технологическое ограничение, наложенное на развитие ИИ и широкую автоматизацию).

Чтобы снизить риски безответственного использования ИИ, правовой статус автономного и неавтономного ИИ должен быть закреплён в законодательстве по-разному. Для полуавтономных ИИ-моделей ответственность за результаты их работы ИИ лежит в той или иной доле на разработчиках ИИ, операторе или контролере. Не регламентирована в законах и ответственность кибернети-

ческих (гибридных) организмов (человек с «киберусилителями» физических или интеллектуальных функций), систем роевого искусственного интеллекта (рои автономных дронов и иных устройств технологии M2M, координирующих свои действия без участия человека в соответствии с меняющейся полевой обстановкой, поэтому существуют проблемы ответственности подобных «коллективных электронных лиц». ИИ используется для создания художественных объектов (генераторы новостей, аналитики; топологий микросхем и программного обеспечения, видео и фото); как использовать подобную «авторскую продукцию»? Автономные технологии являются объектами гражданских прав в той мере, в какой расширяют возможности физического и юридического лица, а автономные единицы ИИ де-факто становятся (хотя пока в России не де-юре) субъектом правовых отношений. У сильного искусственного интеллекта могут возникнуть свои, отличные от человеческих, смыслы, предпочтения, критерии. Поэтому на уровне законодательной базы разработки и использования когнитивных технологий должен быть обеспечен четкий приоритет «человеческого» перед «цифровым» (что не всегда соблюдается на практике и сегодня). В целом нужен не стихийный путь развития искусственных систем, и не тот «узкий» путь, которым пошло общество по инициативе IT-корпораций, заинтересованных в расширении влияния своей продукции, а подконтрольность со стороны общества и государства, с учетом социальных рисков и интересов широкого круга экспертного сообщества и общественности, с привлечением не только технических, но и гуманитарных учетных, специалистов при разработке законопроектов и стандартов, регламентирующих разработку и внедрение инструментов и платформ ИИ, поскольку социальные последствия и риски использования «интеллектуальных технологий» выходят за рамки корпоративных интересов разработчиков или пользователей инструментов ИИ, затрагивая основы существования общества. Первоочередными видятся мероприятия по созданию федерального органа исполнительной власти Российской Федерации, осуществляющего функции по нормативно-правовому регулированию, контролю и надзору в сфере применения ИИ, подготовке законодательства в области созда-

ния и использования ИИ, разработке юридических моделей предупреждения криминального поведения ИИ, в частности определению рисков его применения. Требуется госструктура надзора за развитием технологий искусственного интеллекта, полномочная обеспечивать предупреждение социальных и военных рисков, связанных с внедрением социально-опасных инструментов на основе ИИ.

Список литературы

1. Брадул Н.В. Концептуализация понятия «Smart Government»: наукометрический подход / Н.В. Брадул, Э.М. Лебезова // Управленец. – 2020. – Т.11. – №3. – С. 33–45 [Электронный ресурс]. – Режим доступа: <http://upravlenets.usue.ru/images/85/3.pdf>

2. Визильтер Ю.В. Сильный искусственный интеллект – наследник человечества. Ч. 2. / Визильтер Ю.В., Козлов А [Электронный ресурс]. – Режим доступа: <https://scientificrussia.ru/articles/silnyj-iskusstvennyj-intellekt-naslednik-chelovechestva-chast-2>

3. В Минобороны рассказали о применении искусственного интеллекта в армии [Электронный ресурс]. – Режим доступа: <https://ria.ru/20180315/1516469827.html> (дата публикации: 3.03.2020).

4. Громов А. Как Cambridge Analytica «взламывала выборы» по всему миру / А. Громов [Электронный ресурс]. – Режим доступа: <https://tass.ru/mezhdunarodnaya-panorama/5048632> (дата публикации: 5.04.2018).

5. Искусственный интеллект: аналитический сборник // РАНХиГС. – 2020. – №8 – 43 с. [Электронный ресурс]. – Режим доступа: https://www.dropbox.com/s/gstmjuo5fvxxjoq/%D0%98%D0%BD%D0%B4%D0%B5%D0%BA%D1%81-%D0%98%D0%98-2020_%D0%90%D0%BB%D1%8C%D0%BC%D0%B0%D0%BD%D0%B0%D1%85_%E2%84%968_2021-04-14.pdf?dl=0

6. Касми Э. Россиянина едва не посадили на 8 лет из-за ошибки искусственного интеллекта / Э. Касми // CNews [Электронный ресурс]. – Режим до-

ступа: https://www.cnews.ru/news/top/2020-11-12_rossiyanina_edva_ne_posadili (дата публикации: 12.11.2020).

7. Королев И. Как искусственный интеллект будут внедрять в здравоохранении, транспорте и сельском хозяйстве / И. Королев [Электронный ресурс] / Режим доступа: https://www.cnews.ru/articles/2021-01-25_kak_iskusstvennyj_intellekt_budut (дата публикации: 25.01.2021).

8. Королев И. В России будут раскрывать преступления с помощью искусственного интеллекта / И. Королев [Электронный ресурс]. – Режим доступа: https://www.cnews.ru/articles/2021-02-15_v_rossii_budut_raskryvat_prestupleniya (дата публикации: 15.02.2021).

9. Кузаев М. Особое мнение нейросетей. Как с криминалом борются алгоритмами к ужасу правозащитников / М. Кузаев [Электронный ресурс]. – Режим доступа: <https://tass.ru/obschestvo/5941548>

10. Ларина Е.С. Искусственный интеллект. Большие данные. Преступность / Е.С. Ларина, В.С. Овчинский. – М.: Книжный мир, 2018. – 416 с.

11. Министерства подготовят дорожные карты по внедрению решений на основе искусственного интеллекта. Сайт Правительства России [Электронный ресурс]. – Режим доступа: <https://ac.gov.ru/news/page/ministerstva-podgotovat-doroznye-karty-po-vnedreniu-resenij-na-osnove-iskusstvennogo-intellekta-26685> (дата публикации: 24.08.2020).

12. Минобороны решило сделать тяжелый беспилотник «Охотник» дальним перехватчиком. ТАС [Электронный ресурс]. – Режим доступа: <https://tass.ru/armiya-i-opk/9299951> (дата публикации: 27.08.2020).

13. Носов Н. ЖКХ как арена внедрения интернета вещей / Н. Носов // ИКС. – №3 [Электронный ресурс]. – Режим доступа: <https://www.iksmedia.ru/articles/5669547-ZhKX-kak-arena-vnedreniya-interneta.html> (дата публикации: 5 июня 2020).

14. Овчинский В.С. Этика цифровых технологий в полиции / В.С. Овчинский [и др.] // Аналитический доклад «Этика и «цифра»: этические

проблемы цифровых технологий». – РАНХиГС [Электронный ресурс]. – Режим доступа: https://ethics.cdto.center/7_4

15. Официальный портал эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы [Электронный ресурс]. – Режим доступа: <https://mosmed.ai>

16. Переслегин С. Если мы не сможем контролировать ИИ, то дальше у нас есть два пути... / С. Переслегин, А. Носков [Электронный ресурс]. – Режим доступа: <https://hightech.plus/2019/09/28/esli-mi-ne-smozhem-kontrolirovat-ii-to-dalshe-u-nas-est-dva-puti> (дата публикации: 19.10.2019).

17. Распоряжение Правительства РФ от 19 августа 2020 №2129-р «Об утверждении Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники на период до 2024 г.» [Электронный ресурс]. – Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/74460628>

18. Скрынникова А. Власти определились с искусственным интеллектом в работе чиновников и МЧС / А. Скрынникова, А. Агеева, А. Балашова, П. Кананев [Электронный ресурс]. – Режим доступа: https://www.rbc.ru/technology_and_media/16/12/2020/5fd774869a7947c27f22fe25 (дата публикации: 16.12.2020).

19. Созыкин А.В. Обзор методов обучения глубоких нейронных сетей / А.В. Созыкин // Вестник ЮУрГУ. – Сер. Выч. матем. информ. – 2017. – Т.6. – Вып. 3. – С. 28–59 [Электронный ресурс]. – Режим доступа: <http://www.mathnet.ru/links/645301328891949c4840363cb6c090b2/vyurv170.pdf>

20. Степанян А.И. Предиктивная аналитика в прогностической деятельности полиции современных государств / А.И. Степанян // Вестник Санкт-Петербургского университета МВД России. – 2019. – №4 (84). – С. 43–50 [Электронный ресурс]. – Режим доступа: DOI: 10.35750/2071-8284-2019-4-43-50

21. Указ Президента РФ от 10 октября 2019 №490 «О развитии искусственного интеллекта в Российской Федерации» [Электронный ресурс]. – Режим доступа: <http://publication.pravo.gov.ru/Document/View/0001201910110003>

22. Утверждена Стратегическая карта ФНС России на 2021–2023 годы [Электронный ресурс]. – Режим доступа: https://www.nalog.ru/rn77/news/activities_fts/10675920 (дата публикации: 16 марта 2021).

23. Шмырова В. Искусственный интеллект Amazon закрыли, когда он понял, что женщины хуже мужчин / В. Шмырова [Электронный ресурс]. – Режим доступа: www.cnews.ru/news/top/2018-10-10_iskusstvennyj_intellekt_amazon_ulichili_v_gendernoj (дата публикации: 10.10.2018).

24. Шумский С.А. Нужен ли России искусственный интеллект? // Искусственный интеллект. Аналитический сборник. / С.А. Шумский. – М.: ЦНТИ МФТИ, 2019. – 88 с.

25. AI in Oil and Gas Market – Growth, Trends, and Forecast (2020–2025) [Электронный ресурс]. – Режим доступа: https://www.reportlinker.com/p05881461/AI-in-Oil-and-Gas-Market-Growth-Trends-and-Forecast.html?utm_source=GNW

26. Angwin J. Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks / J. Angwin, J. Larson, S. Mattu, L. Kirchner [Электронный ресурс]. – Режим доступа: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (дата публикации: 23 мая 2016).

27. Artificial Intelligence (AI) in Manufacturing Market Worth \$27 Billion by 2027 // Exclusive Report by Meticulous Research [Электронный ресурс]. – Режим доступа: <https://www.globenewswire.com/news-release/2020/03/30/2008236/0/en/Artificial-Intelligence-AI-in-Manufacturing-Market-Worth-27-Billion-by-2027-Exclusive-Report-by-Meticulous-Research.html>

28. Crime prediction for more agile policing in cities – Rio de Janeiro, Brazil // Case study of U4SSC. 2019 [Электронный ресурс]. – Режим доступа: <https://www.itu.int/en/publications/Documents/tsb/2019-U4SSC-Crime-prediction-for-more-agilepolicing-in-cities-Rio-de-Janeiro-Brazil/index.html#p=1>

29. Heaven, W.D. Google medical AI accurate in a lab / Heaven, W.D. // MIT Technology review [Электронный ресурс]. – Режим доступа: <https://www.technologyreview.com/2020/04/27/1000658/google-medical-ai-accurate-lab-real-life-clinic-covid-diabetes-retina-disease> (дата обращения: 20.01.2021).

30. Rocher L. Estimating the success of re-identifications in incomplete datasets using generative models / Rocher L., Hendrickx J.M., de Montjoye, Y.-A. // Nature Communications. URL: <https://www.nature.com/articles/s41467-019-10933-3> (дата публикации: 23 июля 2019).

Миронова Наталия Геннадьевна – канд. филос. наук, доцент ФГБОУ ВО «Башкирский государственный университет», Россия, Уфа.
