Секция «Искусственный интеллект и «умное» государственное управление: от ретроспективности к проспективности контроля (надзора)»

Угрозы экономической безопасности в сфере технологий искусственного интеллекта

Научный руководитель – Павлов Антон Владимирович

Чернов С.А.[1], *Сикаев Т.А.*[2], *Яунгад Г.Э.*[3], *Федурин А.Н.*[4]

1 - Санкт-Петербургский государственный экономический университет, Факультет бизнеса, таможенного дела и экономической безопасности, Санкт-Петербург, Россия, *E-mail: chernovsa2001@gmail.com*; 2 - Санкт-Петербургский государственный экономический университет, Факультет бизнеса, таможенного дела и экономической безопасности, Санкт-Петербург, Россия, *E-mail: chernovsa2001@gmail.com*; 3 - Санкт-Петербургский государственный экономический университет, Факультет бизнеса, таможенного дела и экономической безопасности, Санкт-Петербург, Россия, *E-mail: chernovsa2001@gmail.com*; 4 - Санкт-Петербургский государственный экономический университет, Факультет бизнеса, таможенного дела и экономической безопасности, Санкт-Петербург, Россия, *E-mail: chernovsa2001@gmail.com*

Один из подходов к определению понятия «искусственный интеллект» отражает данную категорию как область технологий или, скорее, передовую науку, изучающую компьютеры, роботизированные устройства и аналитические системы, которые учат их мыслить так же разумно, как люди [1]. Как известно, идея по созданию «гениального робота-ассистента» зародилась задолго до изобретения первого компьютера.

На наш взгляд, к основным признакам искусственного интеллекта можно отнести:

наличие аналитической системы с интеллектуальным поведением, которая может изучать, прогнозировать и выдвигать гипотезы на основе наборов данных независимо или под индивидуальным наблюдением;

возможность понимания человеческого интеллекта у машин.

По мнению авторов, следует отличать понятия категории «искусственный интеллект» от категории «технологии искусственного интеллекта», которые не обладают собственным сознанием, а лишь являются средствами его достижения.

Безусловно, без технологических принципов невозможно представить развитие искусственного интеллекта (далее - ИИ). Они основаны на моральных законах, обозначенных Айзеком Азимовым в его научно-фантастическом рассказе «Хоровод»:

роботы или системы с искусственным интеллектом не причиняют вреда человеку своими действиями или пассивностью;

робот должен подчиняться приказам, полученным от физических лиц, за исключением приказов, нарушающих закон;

если не нарушены первый и второй законы, робот должен обратить внимание на свою безопасность [2].

Проведенный анализ показал, что на сегодняшний день у ИИ несколько ответвлений развития:

1. Самостоятельное обучение.

Ярким примером является методика «от учителя к учителю». Учителя преподают модели искусственного интеллекта на основе наборов данных, собранных и подписанных людьми в соответствии с заранее определенными категориями.

За последнее десять лет векторное обучение привело к немалому развитию в области технологий, изучающих роботизированные устройства - от автономных машин до голосовых ассистентов, - все из которых имеют ограничения.

Процесс поиска терабайтов данных вручную очень дорог и занимает много времени. Тот факт, что данные должны быть помечены вручную, прежде чем люди смогут обрабатывать модели машинного обучения, стал серьезным препятствием для развития технологий искусственного интеллекта.

С другой стороны, самостоятельное обучение - это метод искусственного интеллекта, в котором алгоритм обучается на основе данных без пометок. Система распознает элементы мироздания на основе остального мира. Анализируя поведение, закономерности и отношения таких источников информации, как фразы в тексте или кадры в видео, система загружает подсказки об окружающей среде.

Наблюдаемая практика более точно отражает то, как люди воспринимают мир посредством открытых исследований и рассуждений.

2. Всестороннее обучение.

Конфиденциальность данных - одна из самых серьезных проблем в эпоху цифровых технологий. Поскольку данные являются важным источником для современного искусственного интеллекта, соображения конфиденциальности данных сыграли важную роль в его развитии.

Представляется, что искусственный интеллект играет все более важную роль в повышении конфиденциальности. Вероятно, самый многообещающий способ защиты конфиденциальности - это всестороннее обучение.

Его суть отражает алгоритм запроса единого набора данных для обучения интегрированной модели, с последующим сохранением данных на периферии для обмена между несколькими машинами и серверами. В результате получается унифицированная модель, которая работает сразу для всего набора данных.

3. Transformers.

Первый Transformer был создан компанией Google в 2017 году и представлял из себя алгоритм, ключевой функцией которого было «внимание» — возможность вычислять вероятность появления того или иного слова среди других слов.

Не так давно компания OpenAI дала старт совершенно новой языковой модели ИИ GPT, которая может писать стихи, работать с функциональным программированием и многое другое. GPT-3 продолжает подход OpenAI, заложенный в GPT и GPT-2. Обе первые версии системы были адаптацией Transformer [3].

Одним из главных нововведений «Трансформеров» было параллельная обработка языков: все буквы в тексте интерпретировались одновременно, а не последовательно. Это позволяет модели видеть взаимосвязь между словами, независимо от того, как далеко они находятся, находить слова и фразы в буквах, требующих внимания.

Представляется, что в ближайшие несколько лет «Трансформеры» заложат основу для нового поколения возможностей искусственного интеллекта, начиная с естественного языка.

В настоящее время активное внедрение технологий, позволяющих имитировать когнитивные функции человека, идет на уровне государственных структур, научно-исследовательских и производственных организаций и на уровне частных компаний.

Вектор развития отечественных технологий в этой сфере определен в Национальной стратегии развития искусственного интеллекта на период до 2030 года, утвержденной Указом Президента Российской Федерации от 10 октября 2019 г. № 490 [4]. Одним из главных основоположников написания стратегии стал президент и председатель правления ПАО «Сбербанк» Герман Греф [5].

Сбербанк является разработчиком «дорожной карты» развития ИИ в России. Именно Сбербанк сыграл роль координатора в создании российской стратегии развития технологий ИИ, которая в значительной степени получилась корпоративной: в написании доку-

мента также участвовали представители «Яндекса», Mail.ru Group и «Газпром нефти».

В рамках федерального проекта «Искусственный интеллект» национальной программы «Цифровая экономика Российской Федерации», а также различных профильных госпрограмм, проектов и «дорожных карт» выработан комплекс соответствующих мер, направленных на реализацию Стратегии [6].

При этом приоритетными направлениями использования технологий искусственного интеллекта в Российской Федерации являются развитие отраслей экономики и социальной сферы.

На наш взгляд, в условиях тотальной цифровизации технологии искусственного интеллекта приобретают форму инструмента геополитического влияния.

У конкурентных стран, к примеру США здесь имеется существенный задел, но это отнюдь не предопределяет их победу в новой гонке. Есть высокие шансы у Китая, который может догнать и перегнать США в этом соревновании. Стратегия Китая опирается на его естественные конкурентные преимущества — лидерство в мобильных платежах и электронной торговле, обилие порождаемых ими данных, высококонкурентную среду малого и среднего бизнеса [7].

Представляется, что стратегия России должна опираться на ее конкурентные преимущества, например на исторически сильную физико-математическую и программистскую школы. Сегодня это реальное преимущество, т.к. центры создания стоимости перемещаются от производства в разработку и дизайн, и основной производительной силой становится интеллект разработчиков.

В связи с этим, по мнению авторов, область искусственного интеллекта неизбежно превращается в сферу не только научно-технической конкуренции, но и военно-политического противоборства.

Учитывая причинообразующие факторы этих угроз, развитые страны, прежде всего США и их ближайшие союзники, на различных международных площадках активно продвигают выгодные им подходы к формированию фундаментальных основ создания систем искусственного интеллекта и регулирования их применения в различных областях.

Очевидно, что целью государств, преуспевших в разработке таких правил и стандартов, и стремящихся к их закреплению на международном уровне, является получение в меняющихся исторических реалиях новых возможностей доминирования в мировом масштабе.

При этом, по мнению авторов, существует риск подмены объективных критериев, касающихся функциональности, надежности и безопасности автоматизированных систем, нечеткими субъективными оценками, которые могут позволить использовать новые технологии для достижения конъюнктурных политических и экономических целей.

Эти обстоятельства формируют угрозы национальной безопасности Российской Федерации и диктуют необходимость адекватного реагирования.

Так, Россия выступает против недобросовестной конкуренции, создания неравных условий для доступа к передовым технологиям и ограничения их продвижения на рынках других стран.

При этом стратегически важной задачей является консолидация усилий всех заинтересованных государственных органов и коммерческих организаций на следующих направлениях:

внедрение российских технологических решений в международные стандарты и другие документы, регулирующие сферу разработки искусственного интеллекта;

продвижение на мировых рынках отечественных разработчиков технологий и продуктов искусственного интеллекта;

обеспечение государственной поддержки процессу подготовки специалистов в сфере

искусственного интеллекта.

Безусловно, с развитием технологий, человеку удается изобретать все более совершенные механизмы и устройства, однако объем всех созданных технических алгоритмов довольно мал в сравнении с объемами человеческих знаний. Другими словами, искусственный интеллект еще не достиг уровня, при котором бы он мог конкурировать с человеческим мозгом. Следовательно влияние ИИ на всемирную экономику на данный момент весьма ограничено. Однако после смены технологического уклада ситуация может в корне измениться. В момент такого перехода всегда появляется «замыкающая» технология, способная спровоцировать новую технологическую революцию. Именно в роли такой «замыкающей» технологии в будущем, с большой долей вероятности, сможет выступить ИИ.

Развитие экономики можно назвать цикличным процессом, в котором в определенный период происходит смена технологических укладов. Каждый уклад характеризуется:

наличием определенного ресурса в большом объеме;

наличием специальных инновационных технологий, использующих этот ресурс;

наличием рынка, подходящего для функционирования в нем таких технологий и ресурсов.

В качестве нового дешевого и эффективного ресурса, на наш взгляд, будут выступать компьютерные мощности, ведь уже в настоящее время мы можем наблюдать стремительное развитие вычислительных сил компьютеров, которые превосходят совокупные вычислительные мощности людей. Более того с течением времени этот ресурс постоянно дешевеет и все больше вовлекается в экономические процессы. Пока что такой серьезный потенциал не сумели раскрыть в полной мере, но с развитием ИИ эта задача становится все более возможной, и именно машинный интеллект в будущем будет способен в полной мере использовать все мощности компьютеров.

Вместе с тем, процесс создания искусственного интеллекта идет крайне медленно, поскольку объем информации устанавливается в память машин лишь небольшим процентом людей, являющимися программистами, и загрузить объем знаний в компьютер, который хотя бы отдаленно сможет приблизиться к объему знаний человечества, является невыполнимой задачей. Но искусственный интеллект может справиться с этой проблемой. Речь идет о функции обучения, ведь именно это мы и наблюдаем в последнее время: системы распознавания речи, машинный перевод, компьютерное зрение и т.д. Человек может, написав определенное число алгоритмов, поручить процесс создания программ самим компьютерам. Таким образом существенно сократятся сроки, за которые машинам удастся достигнуть человеческого уровня развития. При выполнении таких условий ИИ простонапросто не сможет оставаться в стороне и не оказывать существенного влияния на глобальные мировые процессы. Другими словами, именно искусственный интеллект станет основой нового технологического уклада.

Такая тенденция безусловно приведет к возникновению новых угроз, напрямую исходящих от искусственного интеллекта:

1. Угроза использования новых технологий для совершения террористических актов. Важно помнить, что ИИ может применяться не только в благих целях, но и в качестве инновационного оружия. В недалеком будущем террористы смогут освоить автономные технологии и активно применять их в своих целях. Так, например, станет возможным создание автономных дронов, в системе которых будет прописана программа, способная действовать и принимать решения в боевых условиях словно человек. Это лишь пример того, как могут быть задействованы подобные изобретения. Такие технологии существенно увеличат масштаб повреждений и количество человеческих жертв. Это вынудит государственные службы разрабатывать и создавать новые методы борьбы и противодействия подобного рода устройствам.

2. Угроза использования новых технологий для незаконного получения информации. На сегодняшний день компьютерные мощности непрерывно используются злоумышленниками для совершения преступлений. В 2021 году число кибератак увеличилось на 54% в сравнении с 2020 годом и составило в среднем 1550 кибератак на одну организацию в неделю [8]. Несложно представить изменение этой цифры с переходом на новый технологический уклад. Преступники смогут в полной мере реализовать искусственный интеллект используя разные направления. Например:

увеличение мощности и эффективности работы программ-взломщиков;

применение полностью автономных систем, что приведет к колоссальному росту числа кибератак по всему миру;

использование технологий по симуляции голоса в реальном времени;

использование технологий создания фейковых видеоматериалов и т.д.

Такие технологии существуют уже сейчас, однако их использование некоторых из них неэффективно, так как они нуждаются в существенных доработках.

3. Риск уменьшения роли человека в социально-экономических процессах. Некоторые ученые считают, что развитие совершенного искусственного интеллекта станет последним изобретением человечества, говорит руководитель робототехнического центра «Сколково» Альберт Ефимов [9].

Развитие ИИ может уничтожить ряд профессий, для которых ранее требовался человек, ведь уже сейчас роботы заменяют человека в таких видах деятельности как: работник завода, таксист, работник финансовой сферы, переводчик, адвокат и т.д. Также роботы все чаще применяются в быту и выполняют за человека даже самые простейшие действия. В будущем машины смогут выполнять гораздо большее число задач за человека, в результате чего он просто потеряет свои естественные навыки.

Прохождением определенного рубежа в превосходстве машины над человеком может стать момент получения роботами права на легитимное применение насилия [10]. Уже сейчас разработан тест, способный отличить ИИ от человеческого интеллекта, но некоторым компьютерам удается успешно проходить это тестирование.

Исходя из представленных угроз, которые могут быть вызваны развитием технологий искусственного интеллекта, целесообразно разработать рекомендации по их нивелированию.

Для устранения угроз в области создания ИИ необходимо на международном уровне создать нормативно-правовой документ «Международные правила развития и применения искусственного интеллекта». Данный документ предлагается разработать на базе Организации Объединённых Наций. Соответствующий международный акт должен предусматривать сферы применения ИИ, а также основные направления его развития, отображать правила обращения с созданным интеллектом. Данный документ будет содержать основные рекомендации и требования в указанных направлениях.

Для предотвращения угрозы военного применения искусственного интеллекта, а также использования его в террористических актах, документ должен предусматривать полное исключение применения и развития ИИ в военной сфере, запрет на использование ИИ в технических устройствах или объектах военно-промышленного комплекса стран. Эффект от использования искусственного интеллекта в качестве оружия равносилен эффекту от применения оружия массового поражения, поэтому за нарушение вышеуказанного требования акт должен предусматривать наивысшую форму ответственности перед мировым сообществом как за преступление против мира и человечества, которой будут подвергаться как физические лица: разработчики и примирители, так и организации и государства.

Международный акт должен носить превентивный характер. После создания искусственного интеллекта следует разъяснить ему общепринятые правовые нормы, а также

санкции за нарушения прав человека. Для предотвращения угрозы использования новых технологий для незаконного получения информации, а также иных противоправных действий необходимо предусмотреть, что ИИ генерируется не только на территории отдельного государства, но и в международном правовом поле, поэтому искусственный интеллект не должен совершать действий, нарушающих права человека как гражданина не только отдельной страны, но и мира в целом. За несоблюдение искусственным интеллектом требований, в отношении него будут применяться предусмотренные меры, влекущее ограничение его способностей или полное его отключение.

Более того, условием создания ИИ в документе стоит предусмотреть наличие уязвимости перед человеком. Недопустимо создание искусственного интеллекта, который невозможно остановить, отключить или устранить имеющимися на момент его создания силами.

В целях предотвращения угрозы замещения человеческих трудовых ресурсов искусственным интеллектом в документе следует предусмотреть запрет использования ИИ государственными и коммерческими организациями как средство, заменяющей человеческий ресурс, если уровень безработицы по отдельному региону или стране окажется выше пороговых значений. Исключения могут составить только высокотехнологичные отрасли, где невозможно функционирование без применения ИИ. Таким образом, автоматизация должна происходить с учётом возможностей трудоспособного населения сменить место или направление своей деятельности по каждой конкретной отрасли. Если у населения такой возможности не будет, а искусственный интеллект будет просто вымещать человеческие ресурсы, то процесс автоматизации в данной отрасли следует остановить.

Подводя итог, хочется сказать, что в настоящее время многие компании, государства и разработчики направляют много сил на создание искусственного интеллекта. Однако, запуск искусственного интеллекта как новой формы сознания может вызвать ряд угроз не только экономической безопасности стран, организаций, но и безопасности человечества в целом. В настоящий момент имеются НПА, регулирующие развитие технологий ИИ, однако, нет международной принятых документов, регулирующих развитие самого ИИ до и после его создания. Предлагаемый международный нормативно-правовой акт обеспечит нейтрализацию угроз, вызванных созданием ИИ, и станет основой правовых отношений между человеком и искусственным интеллектом, как новой формой сознания.

Источники и литература

1) П. А. Пылов, И. В. Кудаева. / Человек управляет искусственным интеллектом или искусственный интеллект управляет человеком? [Электронный ресурс]// РОССИЯ МОЛОДАЯ Сборник материалов XIII Всероссийской научно-практической конференции с международным участием. — Кемерово: Кузбасский государственный технический университет имени Т.Ф. Горбачева (Кемерово). – 2021. – с. 94703.1-94703.5. URL: https://www.elibrary.ru/item.asp?id=47123934

2) Asimov I. Runaround // Asimov I. I, Robot. — L.: HarperCollins Publishers, 1996. — Pp. 38–60.

3) Нейросеть GPT-3 от OpenAI пишет стихи, музыку и код. Почему она пока далека от настоящего ИИ, но способна поменять мир // NEWS URL: https://news.myseldon.com/ru/

4) Указ Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в российской федерации» // Официальный интернет-портал правовой информации URL: http://publication.pravo.gov.ru/Document/Text/0001201910110003

5) Развитие технологий искусственного интеллекта в России: цели и реальность // Carnegie Endowment for international peace URL: https://carnegieendowment.org/2020/07/07/ru-pub-82173

6) Стратегическое лидерство в цифровую эпоху и технологии искусственного интеллекта // Совет Безопасности Российской Федерации URL: http://www.scrf.gov.ru/news/allnews/3053/

7) Искусственный интеллект: вызовы и угрозы России // Российский совет по международным делам URL: https://russiancouncil.ru/analytics-and-comments/analytics/iskusstvennyy-intellekt-vyzovy-i-ugrozy-rossii/

8) Check Point Research: в 2021 году число кибератак на организации во всем мире выросло на 40% // CRN URL: https://www.crn.ru/news/detail.php?ID=157069

9) Какие угрозы несет человечеству искусственный интеллект // ВЕДОМОСТИ URL: https://www.vedomosti.ru/technology/articles/2016/02/29/631757-kakie-ugrozi-neset-chelovechestvu-iskusstvennii-intellekt

10) Семь смертных грехов искусственного интеллекта // РБК URL: https://trends.rbc.ru/trends/social/5eb299089a79476e9fd77f5c