

Сильный искусственный интеллект: угрозы и перспективы развития в реализации права

В. В. Шадрин, А. И. Спиридонов

СВФУ им. М.К. Аммосова, г. Якутск, Россия

Аннотация. Статья посвящена роли искусственного интеллекта в реализации права, в частности, затрагивается проблема развития «сильного» искусственного интеллекта (англ. Artificial General Intelligence – AGI). В отличие от существующих в настоящее время моделей «слабого» искусственного интеллекта, AGI способен приступить к выполнению тех функций, которые не были изначально предусмотрены, которые требуют общего подхода и обобщенных знаний об окружающем мире в настоящий момент. Авторами исследованы угрозы в применении искусственного интеллекта в настоящий момент и перспективы развития «сильного» искусственного интеллекта в контексте реализации права. Делается вывод о том, что за «сильным» искусственным интеллектом кроется практически потенциал возможностей, благоприятные и негативные последствия его использования зависят от того, каким образом государство и общество распорядятся полученными технологиями.

Ключевые слова: право, реализация права, информационные технологии, цифровизация, Интернет, искусственный интеллект, постиндустриальное общество, «слабый» искусственный интеллект, «сильный» искусственный интеллект, GPT-4.

Artificial General Intelligence: threats and development prospects in the implementation of the law

V. V. Shadrin, A. I. Spiridonov

M.K. Ammosov North-Eastern Federal University, Yakutsk, Russia

Abstract. The article is devoted to the role of artificial intelligence in the implementation of the law, in particular, the problem of the development of “strong” artificial intelligence (English Artificial General Intelligence – AGI) is touched upon. Unlike currently existing models of “weak” artificial intelligence, AGI is able to start performing those functions that were not originally envisaged, requiring a common approach and generalized knowledge about the world around us at the moment. The authors investigated the threats in the use of artificial intelligence at the moment and the prospects for the development of “strong” artificial intelligence in the context of the implementation of the law. It is concluded that behind the “strong” artificial intelligence lies practically the potential of opportunities, the favorable and negative consequences of its use depend on how the state and society dispose of the received technologies.

Keywords: law, implementation of law, digital technologies, digitalization, Internet, artificial intelligence, post-industrial society, “weak” (applied) artificial intelligence, “strong” (general) artificial intelligence, GPT-4.

ШАДРИН Владимир Владиславович – старший преподаватель кафедры «Теория, история государства и права» юридического факультета, Северо-Восточный федеральный университет имени М.К. Аммосова.

E-mail: shadrinworkmail@mail.ru

SHADRIN Vladimir Vladislavovich – senior lecturer of the Department of Theory and History of State and Law, Faculty of Law, M.K. Ammosov North-Eastern Federal University.

СПИРИДОНОВ Алексей Иванович – студент юридического факультета, Северо-Восточный федеральный университет имени М.К. Аммосова.

E-mail: spiridonov_aleksej_iv@mail.ru

SPIRIDONOV Alexey Ivanovich – student, Faculty of Law, M.K. Ammosov North-Eastern Federal University.

Одним из ярких проявлений научно-технического прогресса в наши дни является развитие и применение цифровых, информационных технологий. Один из крупнейших исследователей информационного общества Мануэль Кастельс утверждает, что в постиндустриальном обществе генерирование, обработка и передача информации стали фундаментальными источниками производительности и власти [1, с. 64]. Такое положение дел в равной степени относится и к нашей стране – вступление России в эпоху постиндустриального общества является необходимой реальностью, обусловленной не только развитием интернет-пространства, распространением информации, скоростной передачи данных, но и развитием цифровых технологий [2, с. 14].

В числе прочих продуктов цифровизации существует один особенный – это искусственный интеллект (ИИ). Его можно определить как комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека, комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений [3]. Технологии ИИ предоставляют беспрецедентные возможности для улучшения жизни людей, повышения качества и скорости реализации прав и обязанностей. Внедрение искусственного интеллекта и позитивный результат от его использования, благодаря которому обеспечивается и поддерживается должный уровень комфорта для физических и юридических лиц, желание пользователей применять в жизненных, технологических и иных процессах подобную технику, а также неготовность общества отказаться от нее обуславливают дальнейшее развитие рынка электронных помощников, которое, безусловно, будет связано с увеличением количества умных устройств, а также созданием новых. [4, с. 73].

Однако воздействие подобных технологий на реализацию права может вызывать опасения. Беспокойство по поводу будущих разработок в этой сфере становится всё более обоснованным вследствие возрастающих возможностей существующих систем, их автономности и распространения. Можно привести ряд примеров. Все чаще работодатели государственного и частного секторов обращаются к искусственному интеллекту в целях оптимизировать обработку данных большого количества претендентов на вакантные должности и устранить предвзятость и дискриминацию при приеме на работу [5, с. 74]. Тем самым социальные и экономические права человека ставятся под угрозу. В некоторых контекстах занятости людьми управляют в первую очередь алгоритмические системы принятия решений, что потенциально влияет на их экономические возможности [6, с. 16]. В сфере государственного управления часть административных функций на себя берут камеры видеонаблюдения с применением нейронных сетей, которые отслеживают правонарушения на дорогах, выявляя непристегнутые ремни безопасности у водителей. Это приводит к тому, что административное штрафование переходит на автоматический режим, что несомненно ускоряет реализацию права. Но и здесь пока что не без ошибок или недостатков – искусственный интеллект может воспринять игрушку на переднем сидении автотранспорта как пассажира, который едет без соблюдения необходимых мер безопасности [7]. Таким образом решение ИИ о назначении штрафа как меры наказания (применение права) за совершенный субъектом административный проступок вступает в силу на неправомерной основе. В сфере уголовного права на данный момент можно вспомнить куда более значимый и неблагоприятный прецедент с участием ИИ. Гражданин США Э. Лумис в штате Вискон-

син в 2019 г. был приговорен к шести годам тюремного заключения за угон автомобиля [8]. Как указывает П. Н. Бирюков, прямых улик в ходе рассмотрения дела не хватало, поэтому полиция использовала программу КОМПАС (Correctional Offender Management Profiling for Alternative Sanctions – COMPAS). Программа обосновала свое решение о виновности Э. Лумиса высоким риском рецидива такого рода лица [9, с. 81]. Гражданин не был согласен с тем, что суд использовал решение программы, отчего «справедливое разбирательство» как таковое отсутствовало. Более того, доступ к алгоритму программы был закрыт, однако прокурор заявил, что подобное применение ИИ было в рамках закона и отвечало «научной обоснованности теста». Система, поддерживающая уголовные приговоры с помощью баллов, представляющих риск того, что осужденный преступник совершит дополнительные преступления, должна быть интерпретируемой, проверяемой и открытой для оспаривания обвиняемым, чтобы обеспечить справедливый и открытый судебный процесс [6, с. 15]. Наконец, широкое распространение получило открытое письмо, подписанное значительным количеством пользователей Интернет-сети, в котором выдвигается требование немедленно «остановить гигантский ИИ эксперимент». Как рассуждают авторы открытого письма, системы искусственного интеллекта с человеческим конкурентным интеллектом могут представлять серьезные риски для общества и человечества, как показали обширные исследования и признаны ведущими лабораториями искусственного интеллекта [10].

Приведенные выше примеры являются проявлениями т.н. «слабого» или «прикладного» ИИ (англ. «weak/applied AI»), которые по своей сути являются не более чем инструментом с оптимальным поиском решений поставленной задачи опираясь на свою базу данных и информационные сети. «Слабые ИИ» не могут быть допущены выполнять какие-либо дискреционные полномочия на самостоятельной основе с точки зрения здравого смысла. Однако помимо слабого ИИ мы наблюдаем развитие «сильного» ИИ, чьи потенциальные возможности намного шире и требуют тщательного изучения.

Сильный ИИ на английском языке признано обозначать как искусственный общий интеллект (англ. Artificial General Intelligence). Суть этого названия в том, что сильный ИИ должен уметь решать не одну конкретную задачу и даже не широкий набор задач, он должен быть «человекоподобным» [11, с. 145]. Artificial General Intelligence (AGI) – это представление обобщенных когнитивных способностей человека в программном обеспечении таким образом, чтобы, столкнувшись с незнакомой задачей, система все равно могла найти решение (цель системы AGI состоит в том, чтобы выполнять любую задачу, на которую способен человек) [12]. Следовательно, AGI способен приступить к выполнению тех функций, которые не были изначально предусмотрены и требуют общего подхода, обобщенных знаний об окружающем мире в настоящий момент. Увеличение компьютерных мощностей в скором времени может привести к развитию технологий ИИ, которые, заметим, уже сейчас показывают вполне неплохие результаты во многих сферах общественных отношений и даже способны сдать на приемлемый уровень адвокатские экзамены, включающие в себя не только тестовые вопросы, но и задачи с требованием глубокого анализа ситуации (см. Chat GPT passes the Bar Exam).

В настоящее время технологией, которую можно назвать «искрой AGI», является GPT-4 (англ. Generative Pre-trained Transformer) – это большая языковая модель, разработанная компанией Open AI и является на данный момент наиболее продвинутой версией чат-бота согласно документу Microsoft Research. В документе приведены результаты GPT-4 во многочисленных тестах и испытаниях совершенно разного характера (от написания программного кода и математических вычислений до логических вопросов по расстановке вещей друг на друге). В целом, как замечают сами авторы исследований, во всех задачах

результаты GPT-4 были достаточно близки к человеческому уровню и зачастую превосходили ChatGPT (предыдущую модель чат-бота, работающую на GPT-3.5). Немаловажным стоит отметить факт возможности GPT-4 «взаимодействовать с миром». Пятый раздел в документе от Microsoft Research содержит информацию о том, как данная система ИИ способна выполнять ряд заданий, которые требуют знания действительного окружения, цели и действий пользователя, а также учитывать обратную связь. Таким образом, существует динамическая подстройка дальнейших рекомендаций для пользователя под полученную системой информацию.

Исследователи признают недостатки и ограничения эксперимента, поскольку испытания проходили только с некоторыми проблемами «реального мира» и не могут полностью отражать способности GPT-4 в иных ситуациях [13, с. 53]. Отметим, что технологии с применением ИИ вовсе нельзя называть надежными и совершенными. Они по-прежнему остаются нестабильными и могут время от времени демонстрировать самые различные результаты. Для большего систематического обучения перед машиной необходимо представлять куда более глубокие проблемы нежели условные задачи по расстановке хрупких предметов в пространстве или испытаний по ориентировке в комнатах. К тому же мощность ИИ напрямую зависит от внутреннего устройства системы, характеристик ее составляющей. Все же очевидно, что, повышая вычислительные мощности, рано или поздно удастся добиться в задачах, требующие описания реального мира, результатов, сопоставимых с человеком [11, с. 149].

GPT-4, как и другие языковые модели, обучена на устаревшей информации. Тем не менее данная модель способна использовать внешние инструменты, такие как поисковой движок или программный интерфейс приложений (API) для преодоления многих ограничений [13, с. 43]. С помощью предоставленных инструментов языковая модель способна анализировать предоставленные пользователем изображения, будь то фото произвольного рисунка или написанного от руки текста на простой бумаге. Для правовой сферы это в первую очередь обозначает способность субъекта «общаться» с ИИ. Например, расспрашивать про условия письменного договора, которые самому пользователю были непонятны. Поисковой движок, который использует GPT-4, позволяет компенсировать недостаток знаний ввиду обучения ИИ на устаревшей информации. Предполагаем, что установленная внутри системы база данных в скором будущем может более не являться критической проблемой для применения ИИ в общественных отношениях. Существуют различные информационно-справочные порталы для юристов (например, Официальный интернет-портал правовой информации pravo.gov.ru). Вследствие частых изменений в системе национального и международного законодательства ряд норм утрачивают силу, иные вступают или претерпевают серьёзные изменения в своём содержании. Процесс реализации права весьма сложен, поскольку требует от субъекта ответственности и внимания при ознакомлении с юридическим актом. Потому в сложившихся обстоятельствах способность ИИ получать информацию напрямую с ресурсов с актуальным и действующим законодательством, вероятно, благоприятно скажется на реализации права.

На первый взгляд кажется, что за системами AGI кроется почти безграничный потенциал качественного развития общества. Но в конце концов абсолютно все последствия, как благоприятные, так и негативные, зависят от того, каким образом государство и общество распорядятся полученными технологиями. Всё же с допустимой долей вероятности можно утверждать, что с развитием ИИ роль права в обществе существенно возрастет, а процесс реализации права будет ускорен.

Литература

1. Скрибичский, М. М. Информационная эпоха и новая экономика в трудах Мануэля Кастельса / М. М. Скрибичский // Мир новой экономики. – № 4. – 2015. – С. 62–68.
2. Попова, А. В. Новые субъекты информационного общества и общества знания: к вопросу о нормативном правовом регулировании / А. В. Попова // Журнал российского права. – №11. – 2018. – С. 14–24.
3. Указ Президента РФ от 10.10.2019 г. №490 «О развитии искусственного интеллекта в Российской Федерации» // Собрание законодательства Российской Федерации. 2019 г. № 41. Ст. 5700.
4. Милинчук, Д. С. Искусственный интеллект в праве: тенденции, угрозы и перспективы права на современном этапе развития общества / Д. С. Милинчук // Имущественные отношения в Российской Федерации. – 2021. № 7. – С. 70–76.
5. Алешкова, И. А. Искусственный интеллект и его взаимосвязь с правами человека (Обзор) / И. А. Алешкова // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 4, Государство и право: реферативный журнал. – 2021. – С. 70–78.
6. Leslie, D., Burr, C., Aitken, M., Cows, J., Katell, M., and Briggs, M. (2021). Artificial intelligence, human rights, democracy, and the rule of law: a primer. TheCouncilofEurope.
7. Искусственный интеллект выписал водителю штраф из-за не пристёгнутой игрушки [Электронный ресурс]. – URL: <https://news.ru/regions/iskusstvennyj-intellekt-vypisal-voditelyu-shtraf-iz-za-igrushki/> (дата обращения: 12.08.2023).
8. Loomisv. Wisconsin [Электронный ресурс]. – URL: <https://www.scotusblog.com/case-files/cases/loomis-v-wisconsin> (дата обращения: 12.08.2023).
9. Бирюков, П. Н. Искусственный интеллект и «предсказанное правосудие»: зарубежный опыт / П. Н. Бирюков // Lex Russica. – 2019. – №11. – С. 79–87.
10. Pause Giant AI Experiments: An Open Letter [Электронный ресурс]. – URL: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (дата обращения: 12.08.2023).
11. Редозубов, А. Д. Формализация смысла. Часть 1 / А. Д. Редозубов // Онтология проектирования. – 2021. – №2. – С. 144–153.
12. What is Artificial General Intelligence [Электронный ресурс]. – URL: <https://www.techtarget.com/searchenterpriseai/definition/artificial-general-intelligence-AGI> (дата обращения 18.03.2023)
13. Sparks of Artificial General Intelligence: Early experiments with GPT-4. Microsoft Research. – Искры Искусственного Общего Интеллекта: Ранние эксперименты с GPT-4. Исследования Microsoft [Электронный ресурс]. – URL: <https://arxiv.org/pdf/2303.12712.pdf> (дата обращения: 12.08.2023).

References

1. Skribickij, M. M. Informacionnaja jepoha i novaja jekonomika v trudah Manujelja Kastel'sa / M. M. Skribickij // Mir novoj jekonomiki. – № 4. – 2015. – S. 62–68.
2. Popova, A. V. Novye sub#ekty informacionnogo obshhestva i obshhestva znaniya: k voprosu o normativnom pravovom regulirovanii / A. V. Popova // Zhurnal rossijskogo prava. – №11. – 2018. – S. 14–24.
3. Ukaz Prezidenta RF ot 10.10.2019 g. №490 «O razvitii iskusstvennogo intellekta v Rossijskoj Federacii» // Sbranie zakonodatel'stva Rossijskoj Federacii. 2019 g. № 41. St. 5700.
4. Milinchuk, D. S. Iskusstvennyj intellekt v prave: tendencii, ugrozy i perspektivy prava na sovremennom jetape razvitija obshhestva / D. S. Milinchuk // Imushhestvennye otnoshenija v Rossijskoj Federacii. – 2021. № 7. – S. 70–76.
5. Aleshkova, I. A. Iskusstvennyj intellekt i ego vzaimosvjaz' s pravami cheloveka (Obzor) / I. A. Aleshkova // Social'nye i gumanitarnye nauki. Otechestvennaja i zarubezhnaja literatura. Ser. 4, Gosudarstvo i pravo: referativnyj zhurnal. – 2021. – S. 70–78.
6. Leslie, D., Burr, C., Aitken, M., Cows, J., Katell, M., and Briggs, M. (2021). Artificial intelligence, human rights, democracy, and the rule of law: a primer. TheCouncilofEurope.
7. Iskusstvennyj intellekt vypisal voditelju shtraf iz-za ne pristjognutoj igrushki [Jelektronnyj resurs]. – URL: <https://news.ru/regions/iskusstvennyj-intellekt-vypisal-voditelyu-shtraf-iz-za-igrushki/> (data obrashhenija: 12.08.2023).
8. Loomisv. Wisconsin [Jelektronnyj resurs]. – URL: <https://www.scotusblog.com/case-files/cases/loomis-v-wisconsin> (data obrashhenija: 12.08.2023).

9. Birjukov, P. N. Iskusstvennyj intellekt i «predskazannoe pravosudie»: zarubezhnyj opyt / P. N. Birjukov // LexRussica. – 2019. – №11. – S. 79–87.

10. Pause Giant AI Experiments: An Open Letter [Elektronnyj resurs]. – URL: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (data obrashhenija: 12.08.2023).

11. Redozubov, A. D. Formalizacija smysla. Chast' 1 / A.D. Redkozubov// Ontologija proektirovanija. – 2021. – №2. – S. 144–153.

12. What is Artificial General Intelligence [Elektronnyj resurs]. – URL: <https://www.techtarget.com/searchenterpriseai/definition/artificial-general-intelligence-AGI> (data obrashhenija 18.03.2023)

13. Sparks of Artificial General Intelligence: Early experiments with GPT-4. Microsoft Research. – Iskry Iskusstvennogo Obshhego Intellekta: Rannie jeksperimenty s GPT-4. Issledovanija Microsoft [Elektronnyj resurs]. – URL: <https://arxiv.org/pdf/2303.12712.pdf> (data obrashhenija: 12.08.2023).

