Научная статья
УДК: 347.78
DOI: 10.17323/tis.2024.22306

Original article

CHATGPT, ТЕКСТ, ИНФОРМАЦИЯ: КРИТИЧЕСКИЙ АНАЛИЗ

CHATGPT, TEXT, INFORMATION: CRITICAL ANALYSIS

Марина Николаевна КОМАШКО
Национальный исследовательский университет «Высшая школа экономики», Москва, Российская Федерация,
komashko@list.ru, mkomashko@hse.ru,
ORCID: 0000-0002-5211-0055

Информация об авторе
М.Н. Комашко — доцент департамента права цифровых технологий и биоправа факультета права НИУ ВШЭ, ассоциированный член Кафедры ЮНЕСКО по авторскому праву, смежным, культурным и информационным правам НИУ ВШЭ, кандидат юридических наук

Аннотация. Рассмотрены вопросы теории и практики, связанные с таким типом искусственного интеллекта, как большие языковые модели, в частности ChatGPT. Основное внимание уделено сферам человеческой деятельности, в которых обмен информацией, изложенной в виде текста, имеет наибольшее значение: науке, образованию и журналистике (медиасфере).

Описан опыт взаимодействия пользователей с чат-ботами. Достаточно подробно рассмотрен принцип работы больших языковых моделей. Это позволяет подвести читателя к выводу о том, что чат-боты не могут и не должны осуществлять мыслительный процесс вместо человека и создавать осмысленные, правдивые тексты, которые не нуждались бы в тщательной проверке и редактировании.

Также автор обосновывает вывод о том, что искусственный интеллект (по крайней мере, большие языковые модели) не имитирует деятельность человека, а осуществляет деятельность принципиально иного рода.

В заключительной части работы автор развенчивает миф о том, что чат-боты могут нанести непоправимый ущерб человеческой цивилизации путем внедрения дезинформации.

Ключевые слова: искусственный интеллект, нейросеть, большие языковые модели, LLM, чат-бот, дезинформация, галлюцинирование данными, отравление данными, средства массовой информации, журналистика

Для цитирования: *Комашко М.Н.* ChatGPT, текст, информация: критический анализ // Труды по интеллектуальной собственности (Works on Intellectual Property). 2024. Т. 50, № 3. С. 118–128; DOI: 10.17323/tis.2024.22306

Marina N. KOMASHKO
National Research University Higher School of Economics, Moscow, Russia,
komashko@list.ru, mkomashko@hse.ru,
https://orcid.org/0000-0002-5211-0055

Information about the author
M.N. Komashko — Associate Professor, Department of Digital Law and Bio-Law, Faculty of Law, National Research University Higher School of Economics, Associate Fellow of the UNESCO Chair on Copyright, Neighboring, Cultural and Information Rights at the National Research University Higher School of Economics, Candidate of Legal Sciences

Abstract. The paper deals with theory and practice issues related to such type of artificial intelligence as large language models, in particular, ChatGPT. The main attention is paid to spheres of human activity, in which the exchange of information stated in the form of text is of the greatest importance: science, education and journalism (media sphere).

The experience of user interaction with chatbots is described. The working principle of large language models is discussed in some detail. This allows to lead the reader to the conclusion that chatbots cannot and should not be able to carry out the thinking process instead of a human and create meaningful, truthful texts that would not need careful checking and editing.

The author also substantiates the conclusion that artificial intelligence (at least large language models) does not imitate human activity, but carries out activity of a fundamentally different kind.

In the final part of the paper, the author debunks the myth that chatbots can cause irreparable damage to human civilization by introducing misinformation.

Keywords: artificial intelligence, neural network, large language models, LLM, chatbot, disinformation, data hallucination, data poisoning, media, journalism

For citation: *Komashko M.N.* ChatGPT, Text, Information: Critical Analysis // Trudi po Intellectualnoy Sobstvennosti (Works on Intellectual Property). 2024. Vol. 50 (3). P. 118–128; DOI: 10.17323/tis.2024.22306

В последние годы мы с большим интересом, а порою с обеспокоенностью, наблюдаем за развитием технологий искусственного интеллекта. 2023 г. больше всего, пожалуй, запомнится нам волнениями, связанными с ChatGPT-4. Чего только не происходило в прошлом году в связи с использованием этого программного обеспечения! То студент сдаст выпускную квалификационную работу, написанную чат-ботом, как свою (см., например, [1]; также по этой теме см. [2]), то фанаты решат продолжить фантастическую эпопею вместо ее автора-писателя (см., например, [3]), то ученый заявит в интервью, что «новые инструменты искусственного интеллекта… угрожают выживанию человеческой цивилизации» [4]. Пожалуй, не будет преувеличением сказать, что появление ChatGPT перевернуло нашу жизнь в очередной раз (причем есть ощущение, что этот раз далеко не последний).

Многие пользователи начали экспериментировать и пытаются освоить новый «сорт» искусственного интеллекта. Появились даже курсы и тренинги, на которых обещают научить пользоваться ChatGPT тех, кто не справился сам. В результате экспериментов выяснилось, что ликование (не у всех, кое у кого — ужас) от того, что больше не нужен человек, чтобы состряпать текст на заданную тему, сменилось разочарованием и претензиями то ли к самому искусственному интеллекту, то ли к компании, выпустившей этот продукт на рынок.

Так, коллеги весь год делились по большей части негативными впечатлениями от использования чат-бота и обсуждали проблему выдуманных научных статей, которые никто никогда не публиковал (а в среде юристов также обсуждались выдуманные нейросетью судебные споры). Вот как, например, исследователь Сон Гон Ким описывает свой опыт взаимодействия с ChatGPT: «Я отправил несколько вопросов в ChatGPT … он ответил … я снова спросил: “Есть ли какие-либо ссылки по этой теме?” … он ответил: “Да, … Вот несколько примеров…”. Я проверил, настоящие эти ссылки или поддельные. К сожалению, все ссылки являются поддельными, включая поддельных авторов» [5]. Другие авторы также «столкнулись с тенденцией ChatGPT измышлять факты (феномен галлюцинирования). …Бот сочинял несуществующие цитаты и неточно передавал информацию» [6].

Когда обнаружился такой, мягко говоря, недостаток работы чат-ботов, были предприняты некоторые меры по его нивелированию, но об этом подробнее будет сказано ниже.

Как бы то ни было, способность искусственного интеллекта генерировать связные тексты произвела настолько ошеломляющее впечатление на людей, что возникли представления о способности его также и к анализу текстовой информации. Так, в медицинских кругах проводятся исследования, направленные на определение возможностей использовать чат-бот в диагностике заболеваний. Результаты не слишком удовлетворительные. Например, известно, что «GPT-4 правильно диагностировал только 39% сложных медицинских случаев как у взрослых, так и у детей» [7] (см. также: [8]). Более ранней версии «чат-бот с искусственным интеллектом, работающий на языковой модели под названием GPT-3.5 от OpenAI, не смог правильно диагностировать 83% педиатрических случаев, которые исследовал» [7]. В совокупности «ChatGPT предоставил неправильные диагнозы для 72 из 100 случаев, при этом 11 из 100 результатов были классифицированы как "клинически связанные, но слишком общие, чтобы считаться правильным диагнозом"» [там же].

Тем не менее желание приспособить чат-бот для целей медицины не угасает. В феврале этого года произвела фурор новость об обученном на медицинской литературе GigaChat (аналоге ChatGPT от Сбера), который «сдал» экзамен по направлению подготовки «Лечебное дело» на четыре балла (см., например, [9–11]. Как упоминается в новостных сообщениях об этом событии, и «обучение» искусственного интеллекта, и проверку его способностей проводил Национальный медицинский исследовательский центр им. В.А. Алмазова, что само по себе наводит на определенные мысли (и вызывает вопрос, почему в таких условиях чат-бот сдал экзамен всего на четыре балла, а не на пять с плюсом). Но дело даже не в самом чат-боте и не в нюансах организации проекта. К этой истории мы также вернемся далее.

Журналисты и блогеры не остались в стороне и тоже бросились осваивать ChatGPT. Информационное поле наводнили фейки и курьезы. Примером может служить занятная статейка о последних по состоянию на 2023 г. планах по строительству метро в подмосковной Балашихе [12]. Бесконечно длинная и настолько же бессодержательная заметка полна нелепостей: начиная от знаменитого района Балашихи под названием «Уралмаш» и заканчивая откровением о том, что «в преддверии проведения чемпионата мира по футболу в Балашихе город активно работает над организацией метро для обеспечения удобства перемещения болельщиков и жителей города. В рамках подготовки к чемпионату мира город решил расширить сеть метро» [там же] (в Балашихе метро нет и не было, как и чемпионата мира по футболу. — *Авт.*).

Несмотря на уже всеми, кажется, признанные недостатки текстов от искусственного интеллекта («водянистость», нехватка фактуры, отсутствие эмоционального окраса и смыслов [13], некорректные данные и ошибки [14]), в медиаиндустрии встречается мнение, будто ошибки возникают из-за того, что нейросети «были обучены на неправильных базах данных» [там же], и чат-боты все-таки могут применяться в целях автоматической генерации контента для финансовых новостей, в которых много цифр и фактических данных [там же].

Однако как раз с фактическими данными и возникает главная сложность, что отлично видно на приведенном выше примере. Уже упоминавшийся Сон Гон Ким, в частности, не считает возможным доверять данным от искусственного интеллекта: «…любые новые идеи, сгенерированные ChatGPT, должны быть подтверждены… а результаты должны быть проверены людьми… ChatGPT потенциально может генерировать ложную информацию, поэтому авторам-людям важно тщательно просмотреть и подтвердить информацию, сгенерированную ChatGPT, прежде чем включать ее в свои статьи» [5].

По поводу недостоверности данных, включаемых ChatGPT в свои тексты, высказался также академик А.Р. Хохлов: «Не стоит обвинять в этом нейронную сеть, просто она так работает. Она выдает наиболее вероятный, а вовсе не правильный ответ. Причем наиболее вероятный, исходя из загруженной в нее информации» [15]. Невозможно не согласиться с Алексеем Ремовичем. Будучи человеком с физико-математическим образованием, он обозначил самую суть принципа работы этого программного обеспечения.

Итак, что же такое ChatGPT (кроме того, что он — искусственный интеллект и нейросеть)? ChatGPT относится к семейству больших языковых моделей (Large Language Models, LLM)[1]. У всех LLM общий

---

[1] Необходимо отметить, что в данной статье допускаются не вполне корректные формулировки, когда ставится знак равенства между чат-ботом и искусственным интеллектом. Если говорить более строго, ChatGPT, а также GigaChat и прочие чат-боты — это продукт, выпущенный компанией-разработчиком на рынок, в основе функционирования которого лежит искусственный интеллект, а именно — большая языковая модель. То есть кроме ИИ внутри чат-бота может быть еще много чего, что тоже влияет на качество работы бота. Но поскольку «фундамент» чат-бота составляет именно LLM, а статья предназначена для коллег-гуманитариев (специалисты в компьютерных науках явно будут читать об LLM в других источниках), такая вольность кажется допустимой ради простоты изложения.

принцип работы, и этот принцип заключается в том, что они (как невероятно!) *не пишут тексты* и уж тем более не анализируют и не делают умозаключений. LLM, и ChatGPT в том числе, *составляют цепочку из слов*. По сути, мы все уже давно знакомы с этим механизмом: T9 в смартфоне тоже составляет цепочку из более или менее подходящих, как ему кажется, слов. LLM (ChatGPT) — это более совершенный и эффективный T9. Главная задача чат-бота — продолжить предоставленный пользователем текст на одно следующее слово, а затем продолжить получившийся текст еще на одно слово, затем снова продолжить еще на одно слово и так далее. Как показали эксперименты, наиболее правдоподобные и «читабельные» тексты получаются, если нейросеть настроена продолжать текст не теми словами, которые наиболее часто употребляются после последнего введенного слова, а словами, частота употребления которых следом за «отправным» словом, составляет около 80%.

Замечу, что LLM не понимают значения слов: «…внутри ChatGPT любой фрагмент текста фактически представлен массивом чисел, которые мы можем рассматривать как координаты точки в некоем "пространстве лингвистических признаков". Таким образом, когда ChatGPT продолжает фрагмент текста, это соответствует прослеживанию траектории в пространстве лингвистических признаков» [16]. Естественно, LLM не понимают и значений получающихся текстов. Они выставляют слова по порядку (в соответствии с заложенным в них алгоритмом), но не оценивают смысл, содержание текста и не проверяют достоверность получившихся «данных». Большие языковые модели (как, впрочем, и искусственный интеллект других типов) не знают ничего о мире, они «видят» только предоставленные им для «обучения» тексты, а точнее — знают частоту употребления в этих текстах слов рядом друг с другом.

Есть еще одно важное обстоятельство. LLM обычно запрограммированы на постоянное «дообучение», то есть, взаимодействуя с пользователем или получая доступ к новым текстам в интернете, нейросеть как бы развивает себя. Другими словами, нет какого-то канонического набора текстов, на которые ориентируется языковая модель, она берет в работу все, что попадается «под руку».

Более подробно об устройстве и принципах работы больших языковых моделей можно узнать, например, из [16, 17, 18].

Позволю себе лирическое отступление. Стивен Вольфрам, британский физик, математик и программист, высоко оценивает тексты, генерируемые нейросетями (ChatGPT, в частности): «То, что ChatGPT делает при генерации текста, очень впечатляет — и результаты обычно очень похожи на те, которые получаем мы, люди» [16]. Однако русскоязычные тексты от нейросетей весьма слабы, неинтересны, они «пресные» — сразу видно, что писал их не человек.

Обычно считается, что, раз нейросети разрабатываются по большей части в англоязычных странах, они обучены в основном на текстах, написанных на английском языке, и поэтому лучше им «владеют». Но, не исключено, что объяснение кроется в особенностях самих языков — английского и русского. Возможно, английский как более алгоритмизированный язык, относящийся к группе аналитических языков, лучше поддается «освоению» искусственным интеллектом, чем более стихийный русский, относящийся к группе синтетических языков.

Кроме того, Стивен Вольфрам [16] делает весьма и весьма интересные умозаключения о той роли, которую может сыграть ChatGPT в деле познания законов человеческого мышления и развития человеческого языка, являясь как бы отражением математических закономерностей, существующих в языке, но не обнаруженных пока самими людьми.

Итак, если понять, как устроен и работает ChatGPT, то покажутся странными удивление и возмущение со стороны пользователей из-за того, что нейросеть выдает ложные данные. И еще более странными выглядят надежды на то, что LLM могут что-либо анализировать.

Вероятно, слишком завышенные ожидания от чат-бота возникли из-за того, что изначально кто-то преподнес LLM как программу, способную создавать тексты, подобно человеку. Возможно, в контексте, в котором это было сказано, было совершенно понятно, что именно представляют собой большие языковые модели. Но «мысль изреченная есть ложь» [19], и вырванная из контекста, эта фраза породила неоправдавшиеся надежды.

А ведь никто и не обещал, что чат-бот будет генерировать тексты, содержащие хоть какую-то мысль. Генерация текста ≠ умозаключение. Откуда возьмется мысль или новая идея, если текст — всего лишь достаточно шаблонный набор слов, буквально — общее место [20]? Стоит ли удивляться «водянистости» и «отсутствию смысла» [13]?

Очевидно, что специфическая бессмысленность текстов от чат-бота обусловлена технологией его работы. С другой стороны, людям не всегда нужен какой-либо смысл в тексте. Например, в поздравительных речах не требуется особого смысла. Шаблонные фразы вполне годятся. Вот, например, шутники собрали новогодние речи президента за несколько лет в один поздравительный видеоролик:

(президент появляется в «плиточке» того года, в котором он говорил звучащую в ролике фразу) [21]. Стали ли хуже новогодние поздравления от того, что они повторяются? Очевидно, что нет.

Далее. Никто не обещал, что в текстах от чат-бота будет правда. Если мы понимаем принцип его работы, будем ли мы удивляться, что сведения, сочиненные языковой моделью, не соответствуют действительности? Это ведь тоже следствие технологии.

Никто не обещал, что будет годный контент, который не надо проверять. Как минимум необходимо разбираться в теме, по которой заказал чат-боту текст, и прочитать этот самый текст перед публикацией или отправкой адресату.

Еще один немаловажный момент. В 2023 г. исследователи изучили тенденции саморазвития чат-бота. И — о чудо! — оказалось, что большие языковые модели, постоянно обновляющие базу текстов, со временем (довольно быстро) «тупеют». Этот эффект получил красочные названия: «отравление данными», «крах модели» [22]. Однако, если мы будем помнить о принципе работы и самообучения чат-бота, нам станет очевидно, что со временем «отупение» неизбежно. Как известно, «с кем поведешься, от того и наберешься», а доступные широкой публике языковые модели впитывают сведения и от самих себя (сгенерированные искусственным интеллектом и, как было показано выше, не блещущие качеством), и от пользователей, не все из которых отличаются умом и высокими моральными принципами.

Сама по себе технология больших языковых моделей остается принципиально неизменной, насколько можно судить по доступной информации (имеются сведения о новой технологии, которая потенциально может заменить LLM, но это не относится к предмету данной статьи). Отмеченные выше недостатки в работе чат-ботов нивелируются другими методами.

Так, для устранения проблемы галлюционирования применяется ряд решений (см. об этом [23]). Однако эта проблема все еще остается актуальной (см., например, [24]).

Для борьбы с «водянистостью», нехваткой фактуры и «отравлением данными» тоже нашелся метод — цензура данных для обучения. К примеру, как известно из приведенных выше новостей про GigaChat, его обучали не на всей подряд информации, имеющейся в интернете, без разбора, а только на специализированной медицинской литературе (причем не абы какой, а рекомендованной для обучения будущих врачей). Соответственно подцензурный чат-бот генерирует тексты, опираясь на «веса» слов и словосочетаний в массиве данных, предварительно отобранных разработчиком в соответствии со своей картиной мира. В случае медицины и научного медицинского центра это оправданно; но, как говорится, могут быть и другие варианты.

Теперь, после краткого рассмотрения принципа работы больших языковых моделей, настало время сравнить деятельность человека по написанию текста и деятельность нейросети.

Очевидно, что действия LLM совсем не похожи на то, как человек пишет тексты. Это *принципиально другая* деятельность. Человек, садясь за написание чего-либо, сначала формирует импульс, побуждающий его к этой процедуре, решает «организационные» вопросы: на какую тему он будет писать, в какой форме (письмо, роман, стихи и так далее), с помощью каких инструментов (на бумаге чернилами при свечах или на компьютере); у него в голове рождается мысль, и только после этого человек выражает, более или менее удачно, свою мысль в виде текста. Как бы спонтанно или, наоборот, после раздумий ни взялся живой человек за создание текста, ни при каких обстоятельствах даже самый отъявленный халтурщик от журналистики не станет рассчитывать частоту употребления последовательности слов и составлять слова в цепочку. Человеческий текст — это выраженные вовне эмоции, мысли, идеи (даже если этот человек мыслит штампами, а все его идеи бездарны), но не математические расчеты так называемых «весов».

Говорят, чувственный опыт — самый простой и надежный критерий истины. Поэтому, чтобы окончательно убедиться в том, что LLM составляет тексты *не как человек*, сыграйте в игру: выбирайте случайные слова (существительные, прилагательные, глаголы и так далее) из хорошей книги и собирайте их, одно за другим, в предложения. Текст получится не такой складный, как у чат-бота (что естественно, ведь у этой

игры алгоритм гораздо более примитивный), но он получится! А игроки почувствуют на собственном опыте, насколько вся предыдущая их работа по написанию текстов не похожа на составление цепочек слов (то есть на работу языковой модели).

Также интересно поупражняться в составлении предложений из слов на незнакомом языке (например, китайском или японском, в которых каждый иероглиф, как известно, является целым словом). Если взять иероглифы, относящиеся к одной теме, то при известной доле удачи получится почти настоящий текст. Правда, вы, без знания языка, не поймете, что говорится в составленном вами тексте. Ну так и LLM не понимают смысла составленных предложений, им просто нечем (да и незачем).

И здесь мы подходим к весьма важному моменту. Согласно Национальной стратегии развития искусственного интеллекта, утвержденной Указом Президента Российской Федерации от 10 октября 2019 г. № 490 (ред. от 15 февраля 2024 г.) [25] (далее — Национальная стратегия), искусственный интеллект — это комплекс технологических решений, позволяющий *имитировать* когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их, — комплекс, включающий в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе то, в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений.

Обратимся к толкованию терминов. Имитировать — значит воспроизводить с точностью, подражая кому-нибудь / чему-нибудь [26]. Когнитивные функции человека — это способность понимать, познавать, изучать, осознавать, воспринимать и перерабатывать (запоминать, передавать, использовать) внешнюю информацию [27].

Таким образом, под искусственным интеллектом мы должны понимать комплекс технологических решений, позволяющий воспроизводить с точностью, подражая человеку, способность человека понимать, познавать, изучать, осознавать, воспринимать и перерабатывать внешнюю информацию.

В 2019 г. действительно казалось (и, видимо, кажется до сих пор, раз при внесении изменений в Национальную стратегию определение искусственного интеллекта было слегка изменено, но не было улучшено), что искусственный интеллект имитирует деятельность человека. В те времена чат-ботов еще не было, а нейросети делали такие вещи, как определение объекта, изображенного на фотографии или в кадре видеозаписи, определение соответствия человека фотографии в документе, который этот человек предъявляет, например, в банке для удостоверения личности, считывание номера и тому подобное. Вроде бы это умственная деятельность (физическим трудом такую работу трудно назвать), но все-таки не самая интеллектуальная. Зачастую искусственный интеллект справляется с такими задачами лучше живых людей (хотя результаты бывают разные, в том числе абсолютно курьезные).

Однако, чем интеллектуальнее задача, тем более очевидно, что искусственный интеллект не имитирует работу человеческого мозга. Большие языковые модели осуществляют *совершенно другую* деятельность, ничем не похожую на человеческую, а вовсе не подражают человеку. Причем это утверждение справедливо не только для рассматриваемых здесь LLM, но и для других (хотя, может быть, и не всех) типов искусственного интеллекта (подробнее о некоторых из них см., например, [28, 29]).

Можно с определенными допущениями говорить об имитации *результатов* умственной деятельности человека, да и то — оценки текстов, составленных на разных языках, как было показано выше, существенно разнятся.

Таким образом, мы видим, что определение, данное искусственному интеллекту в Национальной стратегии, основополагающем российском документе об искусственном интеллекте, недостаточно корректно. Оно не соответствует как минимум некоторым типам искусственного интеллекта (в том числе весьма распространенным, таким как большие языковые модели и компьютерное зрение). Есть у этого определения и другие недостатки [28, с. 100].

Другая важная проблема, которую высветил искусственный интеллект, в частности история с GigaChat и экзаменом, касается нас всех как членов общества, но особенно она актуальна для университетов и преподавателей.

Как было показано выше, большие языковые модели не мыслят, ничего не анализируют, ничего не знают о нашем мире, они лишь рассчитывают «веса» и математические траектории и ставят объекты (которые для нас являются словами) в цепочку в соответствии с частотностью их присутствия в неких массивах (для нас являющихся текстами). Из того факта, что LLM справилась с ответами на экзаменационные вопросы, напрашивается следующий вывод: экзамен не проверяет способность студента к анализу и мышлению, *понимание* им темы, а проверяет только его способность вызубрить учебник и ответить близко к тексту (тут у LLM, естественно, большое преимущество). Если выпускнику мединститута не нужны понимание и умение думать, способен ли он лечить людей?

Вот, к примеру, один из экзаменационных вопросов: назначьте пациенту дополнительное обследование [9]. Из самого вопроса студенту уже понятно, что, во-первых, пациент действительно чем-то болеет, а во-вторых, что он недообследован. Дело за малым — назначить процедуры по методичке. А врачу на приеме кто должен подсказать, требуется тут допобследование или нет? Пациент? Или лечащий врач каждый раз к профессору будет за подсказками бегать: что с этим пациентом делать — план лечения составлять или обследовать дополнительно? А если к такому выпускнику на прием попадет здоровый человек, сможет врач его отличить от больного или начнет «лечить анализы»?

Как представляется, вопросы о том, что проверяет экзамен и чему в итоге научены студенты: умению работать по специальности или умению запоминать и пересказывать учебник, касаются не только медицинских учебных заведений, но и всех остальных (просто в них не проводили еще экспериментов с искусственным интеллектом и не хвастались успехами).

Наконец, рассмотрим еще один аспект, который обсуждается в связи с LLM: угроза дезинформации, исходящая от чат-ботов. Доктор Хинтон, которого часто называют крестным отцом искусственного интеллекта, высказал озабоченность в связи с тем, что «интернет будет наводнен фальшивыми фотографиями, видео и текстом, и обычный человек больше не сможет узнать, что является правдой» [30]. Так же считает и Евгений Соколов, специалист в области компьютерных наук и информатики: «Главная угроза лично мне видится в том, что ИИ будет дезинформировать людей и выдавать ложную информацию за истину» [17].

Эту же мысль более развернуто высказал израильский историк и философ Юваль Ноа Харари: «То, что мы обычно принимаем за реальность, часто оказывается просто вымыслом в нашем собственном сознании. … Если мы не будем осторожны (с разработкой и применением искусственного интеллекта. — *Авт.*), мы можем оказаться в ловушке за завесой иллюзий, от которой мы не сможем оторваться — или даже осознать, что она существует» [4].

Действительно, фейки, созданные при помощи ChatGPT, не заставили себя ждать. Уже в мае 2023 г. появились сообщения о задержании полицией Китая человека в связи с подозрением о распространении им в интернете ложных новостей о чрезвычайном происшествии (см., например, [31]). Однако журналисты без всякого чат-бота сами прекрасно справляются с дезинформацией: в частности, сочиняют «кликбейтные» новостные заголовки о том, что в Китае человека арестовали за использование ChatGPT (см., например, [32, 33]).

Дезинформация не является изобретением нейросетей. До появления искусственного интеллекта» дезинформация также существовала, и не просто существовала, а обеспечивала своим адептам славу, почет и даже Пулитцеровскую премию (правда, ненадолго). О нескольких таких случаях, происходивших с 1980-х по 2010-е годы, весьма занимательно рассказал Егор Воробьев [34].

Об этой же проблеме высказался и профессор В.Н. Снетков: «В журналистском творчестве существуют и такие негативные явления, как дифракции этических конвенций, что проявляется в несоответствии действительности публикуемых сведений, девиации установленного порядка их проверки, распространении слухов или версий под видом достоверных сообщений, а в некоторых случаях — в фальсификации общественно значимых фактов, в нарушении принципа полифоничности взглядов на проблему. Это девальвирует способность СМИ осуществлять социальную функцию, гарантировать важнейшее демократическое право граждан на получение достоверной информации» [35, с. 400–401].

Но по большому счету дело даже не в привирающих (иногда очень сильно) журналистах и уж тем более не в чат-ботах, сочиняющих «данные». Это все может быть оставлено на совести отдельных людей, выпустивших негодные тексты в свет.

Существуют более серьезные причины возникновения проблемы дезинформации: «Исследователи акцентируют внимание на отчуждении СМИ от аудитории массовым манипулированием, мифологизацией, мистификацией общественного сознания. Выстраивается жесткая зависимость: собственник — издание — общественное сознание. При уходе из-под контроля СМИ демонстрируют “анархическую независимость, самодостаточность и самовластие, умножая асоциальную прессу и телеканалы, серьезно угрожают свободе массовой информации и гражданским правам в обществе”… Независимость и непонимание в среде СМИ возникают из-за необъятного клубка и пересечения личных интересов, выгод, комплексов, пороков, планов, целей и т.п. … Ряд факторов — экономических, политических, социальных — искажают лицо современной прессы» [там же].

Хотелось бы обратить внимание на то, что все это было исследовано и отмечено еще в 2006 г., когда ни про какой искусственный интеллект среди журналистов никто и не слышал. Да и в более ранние времена существовала проблема качества и достоверности информации. Вот, например, краткая, но емкая характеристика европейских СМИ середины XIX в.: «…характерной особенностью тогдашнего времени была глубокая деморализация прессы. … Никогда еще

не замечалось столь полной и упорной системы подкупа всех органов (прессы. — *Авт.*) большой страны (Франции). Не иначе дело обстоит в Германии. Большинство газет находится в услужении у банков» [36, с. 121–122].

Как видим, вовсе не большие языковые модели и такие продукты современности, как ChatGPT, и даже не искусственный интеллект в целом создают угрозы дезинформации. На качество работы медиаструктур оказывают куда большее влияние совсем другие факторы, которые действовали задолго до появления искусственного интеллекта и продолжат существовать, пока будут существовать люди. Так что в очередной раз мы можем убедиться, что применение искусственного интеллекта не создает принципиально новых проблем, а лишь обращает внимание на давно существующие.

В заключение еще раз кратко сформулирую основные выводы, к которым привело это исследование:

1) прежде чем начинать использовать любые технологии, включая чат-боты, необходимо разобраться в их сути. Да, это непросто, и людям не всегда хочется вникать в особенности технологического процесса, но делать это все равно придется;

2) если есть понимание, как работает технология, то не остается места для появления сюрпризов и для удивления ее «ошибкам» (а на самом деле — абсолютно логичным следствиям применения технологии);

3) способность составить текст не означает умения анализировать, понимать слова, воспринимать информацию;

4) тексты можно создавать разными способами. Человек обычно (если он не играет в предложенные в статье игры) выражает в тексте мысли, идеи, эмоции, образы, которые возникли у него до того, как оказались изложенными в виде текста. То есть первоначальна мысль, текст вторичен. Напротив, LLM составляют цепочки, последовательности слов, одно за другим, не имея представления о том, что из этого получится. Ни получившийся текст, ни даже слова они не воспринимают так, как люди (для них вся «информация», заключенная в словах, — только числа и математические расчеты). Никаких мыслей или эмоций у LLM нет, выражать им нечего. При этом ни отдельные слова, ни итоговый текст не первичны и не вторичны. В «мире» LLM они просто отсутствуют;

5) соответственно LLM, как и некоторые другие виды искусственного интеллекта, *не имитируют деятельность* человека, не подражают ему, а выполняют совершенно другие действия, которые человеку несвойственны (и для человека бессмысленны);

6) с определенными допущениями можно говорить об *имитации результатов деятельности* человека (постольку, поскольку человек признает результат работы LLM в качестве текста);

7) появление LLM высветило некоторые проблемы в сфере образования: если чат-бот успешно справляется с задачей, значит, для ее решения не требуются мышление и аналитические способности. А это, в свою очередь, означает, что необходимо вносить изменения в процессы обучения, чтобы формировать у будущих специалистов понимание, а не запоминание;

8) проблема дезинформации порождена не большими языковыми моделями (чат-ботами) и даже не искусственным интеллектом в целом, а самой природой человеческого общества. Эта проблема вызывала обеспокоенность у мыслителей задолго до появления искусственного интеллекта и обусловлена совершенно другими факторами.

СПИСОК ИСТОЧНИКОВ

1. *Агранович М.* Нейросеть или профессор: дипломную работу за москвича написал искусственный интеллект // Российская газета: сайт. — URL: https://rg.ru/2023/02/01/reg-cfo/nejroset-ili-professor-diplomnuiu-rabotu-za-moskvicha-napisal-iskusstvennyj-intellekt.html (дата обращения: 30.04.2024).
2. *Комашко М.Н.* Экономика, автор и искусственный интеллект: (не)возможности правового регулирования // Право и экономика: стратегии регионального развития: сб. материалов III Вологодского регионального форума с международным участием. Вологда, 22–23 марта 2023 г. Вологда: Северо-Западный институт (филиал) Университета им. О.Е. Кутафина (МГЮА), 2023. С. 28–31.
3. *Дудников К.* Нейросеть ChatGPT дописала серию фэнтези-романов «Песнь льда и пламени» за Джорджа Мартина. — URL: https://www.mentoday.ru/entertainment/news/20-07-2023/neiroset-chatgpt-dopisala-seriyu-fentezi-romanov-pesn-lda-i-plameni-za-djordja-martina/ (дата обращения: 30.04.2024).
4. Yuval Noah Harari argues that AI has hacked the operating system of human civilization // The Economist. 2023. Apr. 28. — URL: https://www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation (дата обращения: 30.04.2024).
5. *Kim S.G.* Using ChatGPT for language editing in scientific articles // Maxillofac Plast Reconstr Surg. 2023. Mar. 8. T. 45(1). P. 13. DOI: 10.1186/s40902-023-00381-x. — URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9992464/ (дата обращения: 30.04.2024).
6. *Голованов Г.* ChatGPT за час написал научную статью с нуля. — URL: https://m.hightech.plus/2023/07/07/

chatgpt-za-chas-napisal-nauchnuyu-statyu-s-nulya (дата обращения: 30.04.2024).
7. *Никифорова А.* Может ли ChatGPT поставить диагноз: ученые провели эксперимент. — URL: https://hightech.fm/2024/01/12/chat-wrong (дата обращения: 30.04.2024).
8. *Barile J., Margolis A., Cason G. et al.* Diagnostic Accuracy of a Large Language Model in Pediatric Case Studies // JAMA Pediatrics. 2024. Jan. 2. DOI: 10.1001/jamapediatrics.2023.5750. — URL: https://jamanetwork.com/journals/jamapediatrics/article-abstract/2813283?resultClick=1 (дата обращения: 30.04.2024).
9. URL: https://sbermed.ai/gigachat-sdal-ekzamen-na-vracha (дата обращения: 30.04.2024).
10. URL: https://lenta.ru/news/2024/02/13/vracha/ (дата обращения: 30.04.2024).
11. URL: https://habr.com/ru/news/793536/ (дата обращения: 30.04.2024).
12. URL: https://investim.guru/obzory/metro-v-balashihe-poslednie-novosti-2023-goda?utm_referrer=https%3A%2F%2Fwww.google.com%2F (дата обращения: 30.04.2024).
13. *Гладкова С.* Нейросеть — конкурент или помощник? — URL: https://journonline.msu.ru/articles/note/neyroset-konkurent-ili-pomoshchnik/ (дата обращения: 30.04.2024).
14. *Арджения Э.* Нейросети и журналистика: как они взаимодействуют. — URL: https://gazeta-ra.info/obshchestvo/item/1008-nejroseti-i-zhurnalistika-kak-oni-vzaimodejstvuyut (дата обращения: 30.04.2024).
15. URL: https://t.me/khokhlovAR/411 (дата обращения: 30.04.2024).
16. *Wolfram St.* What Is ChatGPT Doing ... and Why Does It Work? — URL: https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/ (дата обращения: 30.04.2024).
17. *Соколов Е.* ИИ несет в себе опасности, но совершенно не те, о которых все говорят // URL: https://www.kommersant.ru/doc/6000421 (дата обращения: 30.04.2024).
18. URL: https://youtu.be/VVfFf_XW8zw (дата обращения: 30.04.2024).
19. *Тютчев Ф.И.* Silentium! — URL: https://www.culture.ru/poems/45928/silentium (дата обращения: 30.04.2024).
20. Фундаментальная электронная библиотека «Русская литература и фольклор». — URL: https://feb-web.ru/feb/mas/mas-abc/15/ma257710.htm?cmd=0&istext=1 (дата обращения: 30.04.2024).
21. URL: https://t.me/kremlinlive/4485 (дата обращения: 30.04.2024).
22. URL: https://www.block-chain24.com/articles/snizhenie-proizvoditelnosti-chat-botov-problemy-s-dannymi-ugrozhayut-budushchemu (дата обращения: 30.04.2024).
23. *Миттал А.* Борьба с галлюцинациями в больших языковых моделях: обзор передовых методов. — URL: https://www.unite.ai/ru/%D0%B1%D0%BE%D1%80%D1%8C%D0%B1%D0%B0-%D1%81-%D0%B3%D0%B0%D0%BB%D0%BB%D1%8E%D1%86%D0%B8%D0%BD%D0%B0%D1%86%D0%B8%D1%8F%D0%BC%D0%B8-%D1%81-%D0%BF%D0%BE%D0%BC%D0%BE%D1%89%D1%8C%D1%8E-%D0%B1%D0%BE%D0%BB%D1%8C%D1%88%D0%B8%D1%85-%D1%8F%D0%B7%D1%8B%D0%BA%D0%BE%D0%B2%D1%8B%D1%85-%D0%BC%D0%BE%D0%B4%D0-%B5%D0%BB%D0%B5%D0%B9%2C-%D0%BE%D0%B1%D0%B7%D0%BE%D1%80-%D0%BF%D0%B5%D1%80%D0%B5%D0%B4%D0%BE%D0%B2%D1%8B%D1%85-%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D0%BE%D0%B2/ (дата обращения: 30.04.2024).
24. *Каминский Б.* «Галлюцинации» ChatGPT стали предметом жалобы на конфиденциальность в ЕС // URL: https://forklog.com/news/ai/gallyutsinatsii-chatgpt-stali-predmetom-zhaloby-na-konfidentsialnost-v-es (дата обращения: 30.04.2024).
25. URL: http://static.kremlin.ru/media/events/files/ru/AH4x6HgKWANwVtMOfPDhcbRpvd1HCCsv.pdf (дата обращения: 30.04.2024).
26. *Ушаков Д.Н.* Толковый словарь. — URL: https://dic.academic.ru/dic.nsf/ushakov/823410 (дата обращения: 30.04.2024).
27. *Шмонин А.А.* 5 шагов: как разобраться в когнитивных нарушениях и помочь пациенту. — URL: https://www.1spbgmu.ru/images/home/universitet/Struktura/Kafedry/Kafedra_nevrologii_i_neirohirurgii/Prezentacii_Shmonin/2018/statie/%D0%A8%D0%BC%D0%BE%D0%BD%D0%B8%D0%BD_%D0%90.%D0%90._%D0%9A%D0%BE%D0%B3%D0%BD%D0%B8%D1%82%D0%B8%D0%B2%D0%BD%D1%8B%D0%B5_%D0%BD%D0%B0%D1%80%D1%83%D1%88%D0%B5%D0%BD%D0%B8%D1%8F_-_%D1%80%D0%B5%D0%B0%D0%B1%D0%B8%D0%BB%D0%B8%D1%82%D0%B0%D1%86%D0%B8%D1%8F_%D0%B8_%D0%BB%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_2018.pdf (дата обращения: 30.04.2024).
28. *Комашко М.Н.* Институт авторства и искусственный интеллект // Труды по интеллектуальной собственности. Works on Intellectual Property. 2022. Т. 42. № 3. С. 98–109. — URL: https://doi.org/10.17323/tis.2022.15939

29. *Комашко М.Н.* К вопросу об осмыслении цифровых технологий (на примере искусственного интеллекта) // Цифровое право: технологическая, юридическая и этическая нормативность в условиях функционирования цифровой среды: сб. науч. тр. науч.-аналит. форума: ИФПР. Новосибирск, 2024. С. 117–127.
30. *Metz C.* "The Godfather of A.I." Leaves Google and Warns of Danger Ahead // The New York Times. 2023. May 1. — URL: https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html (дата обращения: 30.04.2024).
31. *Zheng W.* ChatGPT: China detains man for allegedly generating fake train crash news, first known time person held over use of AI bot. — URL: https://www.scmp.com/news/china/politics/article/3219764/china-announces-first-known-chatgpt-arrest-over-alleged-fake-train-crash-news (дата обращения: 30.04.2024).
32. URL: https://digitalbusiness.kz/2023-05-10/v-kitae-arestovali-cheloveka-za-ispolzovanie-chatgpt/ (дата обращения: 30.04.2024).
33. URL: https://dzen.ru/a/ZFpYh8k6vUrSBOBk (дата обращения: 30.04.2024).
34. *Воробьев Е.* Не факт. Четыре невероятные истории журналистов, которые обманули всех // URL: https://disgustingmen.com/history/4-zhurnalista-kotorym-vse-verili-a-zrya/ (дата обращения: 30.04.2024).
35. *Снетков В.Н.* СМИ как законотворческий элемент гражданского общества // Научные труды. Российская академия юридических наук. Вып. 6: в 3-х (4-х) т. Т. 1. М.: Юрист, 2006. С. 399–402.
36. *Петражицкий Л.И.* Акционерная компания. Акционерные злоупотребления и роль акционерных компаний в народном хозяйстве. СПб: Тип. М-ва фин. (В. Киршбаума), 1898.

REFERENCES

1. *Agranovich M.* Neyroset' ili professor: diplomnuyu rabotu za moskvicha napisal iskusstvenniy intellect // Rossiyskaya Gazeta: [sayt]. — URL: https://rg.ru/2023/02/01/reg-cfo/nejroset-ili-professor-diplomnuiu-rabotu-za-moskvicha-napisal-iskusstvennyj-intellekt.html (date of the obrashhenija:30.04.2024 g.).
2. *Komashko M.N.* Ekonomika, avtor i iskusstvenniy intellect: (ne)vozmozhnosti pravovogo regulirovaniya // Pravo i ekonomika: strategii regional'nogo razvitiya: Sbornik materialov III Vologodskogo regional'nogo foruma s mezhdunarodnyim uchastiem, Vologda, 22–23 marta 2023 g. Vologda: Severo-Zapadniy institut (filial) Universiteta imeni O.E. Kutafina (MGYuA), 2023. S. 28–31.
3. *Dudnikov K.* Neyroset' ChatGPT dopisala seriyu fentezi-romanov "Pesn' l'da i plameni" za Dzhordzha Martina. — URL: https://www.mentoday.ru/entertainment/news/20-07-2023/neiroset-chatgpt-dopisala-seriyu-fentezi-romanov-pesn-lda-i-plameni-za-djordja-martina/ (date of the obrashhenija:30.04.2024).
4. Yuval Noah Harari argues that AI has hacked the operating system of human civilization // The Economist. 2023. Apr. 28. — URL: https://www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation (date of the obrashhenija:30.04.2024).
5. *Kim S.G.* Using ChatGPT for language editing in scientific articles // Maxillofac Plast Reconstr Surg. 2023. Mar. 8. T. 45(1). P 13. DOI: 10.1186/s40902-023-00381-x. — URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9992464/ (date of the obrashhenija:30.04.2024).
6. *Golovanov G.* ChatGPT za chas napisal nauchnuyu stat'yu s nulya. — URL: https://m.hightech.plus/2023/07/07/chatgpt-za-chas-napisal-nauchnuyu-statyu-s-nulya (date of the obrashhenija:30.04.2024).
7. *Nikiforova A.* Mozhet li ChatGPT postavit' diagnoz: uchenyie proveli eksperiment // URL: https://hightech.fm/2024/01/12/chat-wrong (date of the obrashhenija:30.04.2024).
8. *Barile J., Margolis A., Cason G. et al.* Diagnostic Accuracy of a Large Language Model in Pediatric Case Studies // JAMA Pediatrics. 2024. Jan. 2. doi:10.1001/jamapediatrics.2023.5750. — URL: https://jamanetwork.com/journals/jamapediatrics/article-abstract/2813283?resultClick=1 (date of the obrashhenija:30.04.2024).
9. URL: https://sbermed.ai/gigachat-sdal-ekzamen-na-vracha (date of the obrashhenija: 30.04.2024).
10. URL: https://lenta.ru/news/2024/02/13/vracha/ (date of the obrashhenija: 30.04.2024).
11. URL: https://habr.com/ru/news/793536/ (date of the obrashhenija:30.04.2024).
12. URL: https://investim.guru/obzory/metro-v-balashihe-poslednie-novosti-2023-goda?utm_referrer=https%3A%2F%2Fwww.google.com%2F (date of the obrashhenija:30.04.2024).
13. *Gladkova S.* Neyroset' — konkurent ili pomoshchnik? — URL: https://journonline.msu.ru/articles/note/neyroset-konkurent-ili-pomoshchnik/ (date of the obrashhenija: 30.04.2024).
14. *Ardzheniya E.* Neyroseti i zhurnalistika: kak oni vzaimodeystvuyut. — URL: https://gazeta-ra.info/obshchestvo/item/1008-nejroseti-i-zhurnalistika-kak-oni-vzaimodejstvuyut (date of the obrashhenija: 30.04.2024).

15. URL: https://t.me/khokhlovAR/411 (date of the obrashhenija: 30.04.2024)
16. *Wolfram St.* What Is ChatGPT Doing ... and Why Does It Work? — URL: https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/ (date of the obrashhenija: 30.04.2024).
17. *Sokolov E.* II neset v sebe opasnosti, no sovershenno ne te, o kotoryikh vse govoryat. — URL: https://www.kommersant.ru/doc/6000421 (date of the obrashhenija: 30.04.2024).
18. URL: https://youtu.be/VVfFf_XW8zw (date of the obrashhenija: 30.04.2024).
19. *Tyutchev F.I.* Silentium! — URL: https://www.culture.ru/poems/45928/silentium (date of the obrashhenija: 30.04.2024).
20. Fundamental'naya electronnaya biblioteka "Russkaya literatura i fol'klor". —URL: https://feb-web.ru/feb/mas/mas-abc/15/ma257710.htm?cmd=0&istext=1 (date of the obrashhenija:30.04.2024).
21. URL: https://t.me/kremlinlive/4485 (date of the obrashhenija:30.04.2024).
22. URL: https://www.block-chain24.com/articles/snizhenie-proizvoditelnosti-chat-botov-problemy-s-dannymi-ugrozhayut-budushchemu (date of the obrashhenija: 30.04.2024).
23. *Mittal A.* Bor'ba s gallutsinatsiyami v bol'shikh yazykovyh modelyah: obzor peredovykh metodov. — URL: https://www.unite.ai/ru/%D0%B1%D0%BE%D1%80%D1%8C%D0%B1%D0%B0-%D1%81-%D0%B3%D0%B0%D0%BB%D0%BB%D1%8E%D1%86%D0%B8%D0%BD%D0%B0%D1%86%D0%B8%D1%8F%D0%BC%D0%B8-%D1%81-%D0%BF%D0%BE%D0%BC%D0%BE%D1%89%D1%8C%D1%8E-%D0%B1%D0%BE%D0%BB%D1%8C%D1%88%D0%B8%D1%85-%D1%8F%D0%B7%D1%8B%D0%BA%D0%BE%D0%B2%D1%8B%D1%85-%D0%BC%D0%BE%D0%B4%D0-%B5%D0%BB%D0%B5%D0%B9%2C-%D0%BE%D0%B1%D0%B7%D0%BE%D1%80-%D0%BF%D0%B5%D1%80%D0%B5%D0%B4%D0%BE%D0%B2%D1% 8B%D1%85-%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D0%BE%D0%B2/ (date of the obrashhenija:30.04.2024).
24. *Kaminskiy B.* "Gallutsinatsii" ChatGPT stali predmetom zhaloby na konfidentsial'nost' v ES. — URL: https://forklog.com/news/ai/gallyutsinatsii-chatgpt-stali-predmetom-zhaloby-na-konfidentsialnost-v-es (date of obrashhenija:30.04.2024).
25. URL: http://static.kremlin.ru/media/events/files/ru/AH4x6HgKWANwVtMOfPDhcbRpvd1HCCsv.pdf (date of the obrashhenija:30.04.2024).
26. *Ushakov D.N.* Tolkovyiy slovar'. — URL: https://dic.academic.ru/dic.nsf/ushakov/823410 (date of the obrashhenija:30.04.2024).
27. *Shmonin A.A.* 5 shagov: kak razobrat'sya v kognitivnyikh narusheniyakh i pomoch' patsientu. — URL: https://www.1spbgmu.ru/images/home/universitet/Struktura/Kafedry/Kafedra_nevrologii_i_neirohirurgii/Prezentacii_Shmonin/2018/statie/%D0%A8%D0%BC%D0%BE%D0%BD%D0%B8%D0%BD_%D0%90.%D0%90._%D0%9A%D0%BE%D0%B3%D0%BD%D0%B8%D1%82%D0%B8%D0%B2%D0%BD%D1%8B%D0%B5_%D0%BD%D0%B0%D1%80%D1%83%D1%88%D0%B5%D0%BD%D0%B8%D1%8F_-_%D1%80%D0%B5%D0%B0%D0%B1%D0%B8%D0%BB%D0%B8%D1%82%D0%B0%D1%86%D0%B8%D1%8F_%D0%B8_%D0%BB%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_2018.pdf (date of the obrashhenija: 30.04.2024).
28. *Komashko M.N.* Institut avtorstva i iskusstvenniy intellect // Trudyi po intellectual'noy sobstvennosti. Works on Intellectual Property. 2022. T. 42. No 3. S. 98–109. — URL: https://doi.org/10.17323/tis.2022.15939
29. *Komashko M.N.* K voprosu ob osmyislenii tsifrovyikh tekhnologiy (na primere iskusstvennogo intellecta) // Tsifrovoe parvo: tekhnologicheskaya, yuridicheskaya i eticheskaya normativnost' v usloviyakh funktsionirovaniya tsifrovoy sredyi: sb. nauch. tr. nauch.-analit. foruma: IFPR. Novosibirsk, 2024. S. 117–127.
30. *Metz C.* "The Godfather of A.I." Leaves Google and Warns of Danger Ahead // The New York Times. 2023. May 1. — URL: https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html (date of the obrashhenija:30.04.2024).
31. *Zheng W.* ChatGPT: China detains man for allegedly generating fake train crash news, first known time person held over use of AI bot. — URL: https://www.scmp.com/news/china/politics/article/3219764/china-announces-first-known-chatgpt-arrest-over-alleged-fake-train-crash-news (date of the obrashhenija: 30.04.2024).
32. URL: https://digitalbusiness.kz/2023-05-10/v-kitae-arestovali-cheloveka-za-ispolzovanie-chatgpt/ (date of the application: 30.04.2024).
33. URL: https://dzen.ru/a/ZFpYh8k6vUrSBOBk (date of the obrashhenija: 30.04.2024).
34. *Vorobyev E.* Ne fact. Chetyire neveroyatnyie istorii zhurnalistov, kotoryie obmanuli vsekh. — URL: https://disgustingmen.com/history/4-zhurnalista-kotorym-vse-verili-a-zrya/ (date of the obrashhenija:30.04.2024).
35. *Snetkov V.N.* SMI kak zakonotvorcheskiy element grazhdanskogo obshchestva // Nauchnyie trudyi. Rossiyskaya akademiya yuridicheskikh nauk. Vyip. 6. V trekh (chetyirekh) t. T. 1. M.: Yurist, 2006. S. 399–402.
36. *Petrazhitskiy L.I.* Aktsionernaya kompaniya. Aktsionernyie zloupotrebleniya i rol' aktsionernyikh kompaniy v narodnom khozyaystve. SPb., 1898.

Научная статья
УДК: 347.78
DOI: 10.17323/tis.2024.22308

Original article

КРИТЕРИИ ТВОРЧЕСТВА ДЛЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

CREATIVITY CRITERIA FOR ARTIFICIAL INTELLIGENCE

Сона Вагановна ДЕ АПРО
Московский государственный университет им. М.В. Ломоносова, Москва, Российская Федерация,
s.deapro@yandex.ru,
ORCID: 0009-0005-6056-3408

Информация об авторе
С.В. Де Апро — аспирант кафедры компьютерного права и информационной безопасности Высшей школы государственного аудита МГУ им. М.В. Ломоносова

Аннотация. В доктрине и научной литературе до сих пор не выработаны универсальные критерии творчества, которые могли бы стать надежными ориентирами в предоставлении защиты прав на результаты интеллектуальной деятельности. Вопрос об авторстве и признании «творчества» искусственного интеллекта также остается открытым. В статье рассмотрена проблема творчества с юридической точки зрения, в частности вопросы охраноспособности произведений, созданных с использованием искусственного интеллекта и непосредственно искусственным интеллектом. Обсуждаются роль воображения в творчестве и другие субъективные факторы, влияющие на творческую деятельность. Изложен взгляд на «мышление» искусственного интеллекта как алгоритмическое, лишенное человеческого духа. Обосновано предложение признать произведения, «созданные» искусственным интеллектом, вторичными, лишенными новизны объектами.

Ключевые слова: творчество, критерии творчества, искусственный интеллект, ИИ, объекты, создаваемые ИИ, результат творческой деятельности, алгоритмы, авторское право

Для цитирования: *Де Апро С.В.* Критерии творчества для искусственного интеллекта // Труды по интеллектуальной собственности (Works on Intellectual Property). 2024 Т. 50, № 3. С. 129–134; DOI: 10.17323/tis.2024.22308

Sona V. DE APRO
Lomonosov Moscow State University, Moscow, Russian Federation,
s.deapro@yandex.ru,
ORCID: 0009-0005-6056-3408

Information about the author
S.V. De Apro — postgraduate student of the Department of Computer Law and Information Security at the Higher School of State Audit, Lomonosov Moscow State University

Abstract. The doctrine and scientific literature have not yet developed universal criteria for creativity that could become reliable guidelines in providing protection for rights to the results of intellectual activity. The question of authorship and recognition of AI "creativity" also remains open. The problem of creativity is considered from a legal point of view, in particular, the protectability of works created using artificial intelligence and directly by artificial intelligence. The role of imagination in creativity and other subjective factors influencing creative activity are discussed. It is proposed to consider the "thinking" of artificial intelligence as algorithmic, devoid of the human spirit. The proposal is substantiated to recognize works "created" by artificial intelligence as secondary objects devoid of novelty.

Keywords: creativity, creativity criteria, artificial intelligence, objects created by AI, the result of creative activity, algorithms, copyright

For citation: *De Apro S.V.* Creativity Criteria for Artificial Intelligence // Trudi po Intellectualnoy Sobstvennosti (Works on Intellectual Property). 2024 Vol. 50 (3) P. 129–134; DOI: 10.17323/tis.2024.22308