

ФИЛОСОФСКИЕ ВОПРОСЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ДОВЕРИЯ «УМНЫМ» АЛГОРИТМАМ В МЕДИЦИНЕ

Т. С. Георгиу ✉

Московский государственный областной университет, Москва, Россия

В статье рассматриваются философские вопросы, связанные с внедрением систем искусственного интеллекта (ИИ) в медицину. В настоящее время использование технологий ИИ является одним из важнейших трендов мирового здравоохранения. «Умные» системы ИИ способны обучаться на собственном опыте, адаптироваться в окружающей среде и принимать решения, которые раньше мог принимать только человек. Эти технологии дают возможность поднять диагностику, лечение и систему профилактики заболеваний на новый уровень. Особое внимание уделено этике и моральным обязательствам разработчиков и пользователей — медработников при переходе к цифровой медицине.

Ключевые слова: искусственный интеллект, машинное обучение, «умные» системы, медицина, биоэтика, моральные решения, доверие, безопасность

Вклад авторов: статья является частью исследовательской работы диссертации автора по теме «Искусственный интеллект: социальные риски».

✉ **Для корреспонденции:** Takis Sophokli Georgiou
Exegetseos, 4, Perivolia, Larnaca, 7560 Cyprus; takis.georgiou.s@gmail.com

Поступила: 01.08.2021 **Статья принята к печати:** 20.08.2021 **Опубликована онлайн:** 30.09.2021

DOI: 10.24075/medet.2021.019

PHILOSOPHICAL ISSUES OF ARTIFICIAL INTELLIGENCE AND “SMART” ALGORITHMS’ TRUST IN MEDICINE

Georgiou TS ✉

Moscow State Regional University, Moscow, Russia

In this article, the author examines the philosophical issues associated with the introduction of artificial intelligence (AI) systems in medicine. Currently, the use of AI technologies in the field of medical sciences is one of the most important trends in world of health. “Smart” AI systems are able to learn from their own experience, adapt in the environment and, according to the parameters of the assigned tasks, can make decisions that in the past belonged only to humans. These AI technologies provide an opportunity to take diagnostics, treatment and disease prevention to a higher level. Particular attention is paid to the ethics and moral obligations of AI developers and healthcare professionals in the transition to such digital medicine.

Keywords: artificial intelligence, machine learning, «smart» systems, medicine, bioethics, moral decisions, trust, safety

Author contribution: The article is part of the research work of the author's dissertation on the topic “Artificial Intelligence: Social Risks”.

✉ **Correspondence should be addressed:** Takis Sophokli Georgiou
Exegetseos, 4, Perivolia, Larnaca, 7560 Cyprus; takis.georgiou.s@gmail.com

Received: 01.08.2021 **Accepted:** 20.08.2021 **Published online:** 30.09.2021

DOI: 10.24075/medet.2021.019

В современном обществе искусственный интеллект (ИИ) охватывает всю повседневную жизнь человека. Это связано с тем, что ИИ является важной составляющей продуктов и устройств повседневной жизни человека. Системы ИИ становятся средой и участниками социальных взаимодействий [1]. Внедрение систем ИИ в медицину — это один из важнейших трендов современного общества. Технологии ИИ существенно меняют мировую систему здравоохранения, формируют революционную систему медицинской диагностики, разрабатывают новые лекарственные средства и в целом способствуют повышению качества медицинских услуг.

В научной литературе даются разные определения термина «искусственный интеллект». По мнению некоторых ученых [2], ИИ представляет собой группу разработанных и закодированных человеком рационально-логических, формализованных правил, которые позволяют имитировать интеллектуальные структуры, воспроизводить целерациональные действия, а также осуществлять последующее кодирование и принятие инструментальных решений без заранее заданного алгоритма. То есть интеллектуальные системы, которые на рынке называются «умными» системами, имеют возможность действовать

автономно. ИИ отличается от традиционных компьютерных алгоритмов тем, что он имеет способность самообучения на основе накопленного «опыта». Эта уникальная функция позволяет ИИ действовать по-разному в аналогичных ситуациях в зависимости от раннего «опыта». «Умные» системы с ИИ обладают рядом характеристик: «интеллектуальность», «разумность», «рациональность», «возможность мыслить как человек» во всех или определенных обстоятельствах.

Конечно, такое определение термина ИИ вызывает споры среди ученых. Некоторые исследователи утверждают, что невозможно моделировать все функции человеческого мозга, а другие уверены, что ИИ может даже превосходить интеллект человека. Также необходимо отметить, что существуют два основных вида ИИ: слабый (англ. «narrow») и сильный (англ. «general», «strong») ИИ [3]. Под «сильным» ИИ понимается способность интеллектуальной системы мыслить и осознавать себя как отдельного субъекта и, в частности, понимать собственные мысли. Такой мыслительный процесс будет подобен человеческому. Существует другой вид ИИ — «слабый ИИ», который обозначает все системы, помогающие человеку решать задачи и достигать поставленные цели.

Вмешательство самого человека здесь необходимо, и приложения подобного рода уже вошли в нашу жизнь: мы живем в окружении техники со «слабым» ИИ и являемся ее постоянными пользователями. Эти технологии вызывают огромный интерес и в медицине, их широкое применение можно найти в «умных» системах здравоохранения.

Внимание к ИИ уделяется и со стороны правительства РФ. Так, согласно Указу Президента Российской Федерации «О развитии искусственного интеллекта в Российской Федерации», «...создание универсального (сильного) искусственного интеллекта, способного, подобно человеку, решать различные задачи, мыслить, взаимодействовать и адаптироваться к изменяющимся условиям, является сложной научно-технической проблемой, решение которой находится на пересечении различных сфер научного знания — естественно-научной, технической и социально-гуманитарной. Решение этой проблемы может привести не только к позитивным изменениям в ключевых сферах жизнедеятельности, но и к негативным последствиям, вызванным социальными и технологическими изменениями, которые сопутствуют развитию технологий искусственного интеллекта» [4]. В соответствии с вышеупомянутым Указом Президента, использование технологий ИИ в социальной сфере должно способствовать созданию условий для улучшения уровня жизни населения, в первую очередь за счет повышения качества услуг в сфере здравоохранения [4].

В современной системе здравоохранения использование ИИ считается одним из важных элементов при предоставлении медицинских услуг. Для достижения целей современной медицины существуют такие технологии, как интеллектуальный анализ данных, экспертные системы, нейронные сети, эволюционные алгоритмы и биокомпьютинг. С использованием ИИ связаны большие надежды. ИИ может сделать здравоохранение более эффективным и удобным для пациентов, ускорить диагностику и уменьшить количество ошибок медицинского обслуживания, помочь пациентам справиться с симптомами или с хроническими заболеваниями, помочь избежать человеческих предубеждений и ошибок. Но использование таких инструментов, предполагает накопление и анализ большого объема биологических данных миллионов больных, которые должны быть сопоставлены с клиническими данными. Экспертные системы являются одной из технологий ИИ, которые применяются для решения неформализованных задач. Отметим, что для этих задач не существует однозначного алгоритмического подхода.

В настоящее время разработаны такие амбициозные проекты, как IBM Medical Sieve, суперкомпьютер IBM Watson, AI-RAD Companion Brain MR for Morphometry Analysis, AI-RAD Companion Prostate MR for Biopsy Support, российская программа на базе технологий ИИ «сервис Цельс», всемирный проект благосостояния World Well-BEING PROJECT и др. Эти проекты нацелены на создание «умных помощников» с многоуровневыми аналитическими способностями, которые будут иметь доступ к накопленным в клинической практике знаниям и способность «мыслить» таким образом, чтобы это помогло принимать решения в различных областях медицины.

Естественно, использование больших данных и технологий ИИ позволяет поднять на новый уровень диагностику, лечение и систему профилактики заболеваний. Такое применение ИИ в медицине, которое требует доверия к «умным системам», актуализирует

фундаментальные философские вопросы. При переходе к цифровой медицине проблемы этики стали ключевыми и во многом определяют скорость технологического прогресса в этой сфере [5]. Исходя из этого, остро стоит вопрос, в какой мере можно использовать данные о здоровье граждан для обучения ИИ и, следовательно, до какой степени можно доверять «умным» системам.

Основная характеристика «умных» систем — это самообучение. Однако следует отметить, что существуют серьезные риски, связанные с правильным самообучением и с проблемой определения его границ. Переобучение является главной угрозой для эффективного лечения. Заметим, что «умные» системы сами по себе не являются умными. Они построены на разных технологиях ИИ, в зависимости от поставленных задач с целью эффективной реализации требуемых функций на основе разработанных алгоритмов. Эти технологии разрабатываются специалистами в области ИИ.

Поэтому Наффилдский совет (Великобритания) [6] — независимая организация, оценивающая множество этических проблем в области биологии и медицины — отмечает, что среди этих вопросов существуют такие, которые следует учитывать в первую очередь.

- Кто отвечает за решения, принимаемые системами ИИ?
- Приведет ли растущее использование ИИ к потере контакта с людьми при оказании медицинской помощи?
- Что произойдет, если системы ИИ будут взломаны?
- Как можно доверять системам, которые в любое время могут оказаться неуправляемыми?

Кроме того, Наффилдский совет не исключает потенциальной возможности принятия ошибочных решений, и в таких случаях возникает вопрос о том, кто несет ответственность, когда ИИ используется для принятия решений [6]. Действия алгоритмов машинного обучения непрозрачны и не дают людям возможности понять, почему ИИ приходит к тем или иным ассоциациям или выводам, и неизвестно, когда может произойти сбой. Следовательно, проверка результатов систем ИИ является сложной проблемой.

В медицине не существует абсолютных знаний, но есть различные подходы к объекту познания. Следовательно, бывают разные результаты и выводы. Поэтому возникает проблема с используемыми данными, применяемыми для обучения систем ИИ. Кроме того, доверие «умным» системам требует обеспечения защиты данных, особенно когда речь идет о конфиденциальных сведениях.

Известный австрийский математик Курт Гёдель в 1931 г. выяснил: любая формальная система, в том числе сама математика, является неполной или противоречивой. Это означает, что существуют проблемы, утверждения и вопросы, которые нельзя ни решить, ни доказать, ни опровергнуть, не выходя за их пределы [7]. Всем известно, что математика является основой для алгоритмов ИИ. Если все детерминировано, тогда машина не может решать свободно новые и неопределенные задачи. А если предоставить машине решать самостоятельно, тогда из вышесказанного следует, что существуют определенные ситуации, в которых машина проявит непредсказуемые реакции. В обоих случаях будут возникать существенные проблемы. Таким образом, любая «умная» система не может самостоятельно принимать решения без вмешательства человека, особенно когда речь идет о моральных поступках [8] для медицинских целей.

Некоторые ученые утверждают, что машинное обучение является хорошим инструментом для совершенствования ИИ. Однако трудно объяснить, как это все работает, и, таким образом, алгоритмы оказываются загадочными даже для их создателей. Неспособность людей понимать данную технологию подрывает доверие к системам ИИ. Доверие имеет решающее значение и является предпосылкой для одобрения в обществе. Многие ученые и разработчики опасаются рисков, если ИИ будет развиваться без наблюдения и руководства со стороны человека. Эти два аспекта — доверие и контроль — являются основой для конструирования надежного и безопасного ИИ [9]. Люди доверяют «умным» системам самое важное: деньги, здоровье и безопасность. Это значит, что мы уже не просто пользуемся технологиями, а зависим от них. Это делает нас уязвимыми.

Этика ИИ определяет моральные обязательства и обязанности разработчиков и пользователей. В свою очередь, биоэтика касается этических вопросов, которые могут возникать при проектировании и разработке систем ИИ в медицине. Следовательно, без согласования важных

аспектов развития ИИ с позиций биоэтики под угрозой могут оказаться медицина и общество в целом [10, 11].

Gartner — самая известная исследовательская и консалтинговая компания, специализирующаяся на рынках информационных технологий — часто публикует отчеты о тенденциях развития технологий. Специалисты компании полагают, что практически все технологии, которые в ближайшее десятилетие будут оказывать существенное влияние на бизнес, людей и общество, связаны с ИИ [12]. Внедрение технологий ИИ — это сложный процесс, который постепенно будет охватывать все параметры личной, профессиональной и социальной деятельности человека, в том числе его здоровье.

Таким образом, технологии ИИ в медицине могут использоваться во благо человека, но абсолютное доверие к «умным» алгоритмам может стать источником множества рисков и опасностей. Поэтому на ученых, разработчиков, инвесторов и других заинтересованных социальных акторов возложена огромная ответственность обеспечить безопасность человека и общества, предусматривая явные и латентные риски, учитывая нормы и принципы биоэтики.

Литература

1. Резаев А. В., Трегубова Н. Д. «Искусственный интеллект», «онлайн-культура», «искусственная социальность»: определение понятий «искусственный». Мониторинг общественного мнения: Экономические и социальные перемены. 2019; 6 (154): 35–47.
2. Пройдаков Э. М. Современное состояние искусственного интеллекта. Цифровая экономика. 2018; 3 (3): 129–153.
3. Searle JR. Minds, brains, and programs. Behavioral and brain sciences. 1980; 3 (3): 417–457.
4. Указ Президента Российской Федерации от 10.10.2019 № 490 [Электронный ресурс]. 2019 [дата обращения 04.04.2021].
5. Доступно с: <http://www.kremlin.ru/acts/bank/44731>
6. Гусев А. В. Этика цифровой медицины [Электронный ресурс]. 2021 [дата обращения 27.07.2021]. Доступно с: https://ethics.cdto.center/7_2
7. Nuffield Council on Bioethics. Artificial intelligence (AI) in healthcare and research. London: Nuffield Council on Bioethics. 2018; 8 p.
8. Gödel K. On Formally Undecidable Propositions of Principia Mathematica and Related Systems I (1931). New York: Dover Publications Inc. 1992; 79 p.
9. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, et al. AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines. 2018; (28): 689–707.
10. Siau K, Wang W. Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI. Journal of Database Management. 2020; 2 (37): 74–87.
11. Кожевникова М., Карпова С. В. Искусственный интеллект: субъект и объект. Этнографическое обозрение. 2020; (1): 80–94.
12. Anderson SL, Anderson M. AI and ethics. AI and Ethics. 2021; 1 (1): 27–31.
13. Panetta K. Trends Drive the Gartner Hype Cycle for Emerging Technologies [Internet]. 2020 [cited 2021 April 05]. Available from: <https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020/>

References

1. Rezaev AV, Tregubova ND. «Iskusstvennyj intellekt», «onlajn-kul'tura», «iskusstvennaja social'nost'»: opredelenie ponjatij «iskusstvennyj». Monitoring obshhestvennogo mnenija: Jekonomicheskie i social'nye peremeny. 2019; 6 (154): 35–47. Russian.
2. Projdakov Je M. Sovremennoe sostojanie iskusstvennogo intellekta. Cifrovaja jekonomika. 2018; 3 (3): 129–153. Russian.
3. Searle JR. Minds, brains, and programs. Behavioral and brain sciences. 1980; 3 (3): 417–457.
4. Ukaz Prezidenta Rossijskoj Federacii ot 10.10.2019 № 490 [Internet]. 2019 [cited 2021 April 04]. Available from: <http://www.kremlin.ru/acts/bank/44731> Russian.
5. Gusev AV. Jetika cifrovov mediciny [Internet]. [cited 2021 July 07]. Available from: https://ethics.cdto.center/7_2. Russian.
6. Nuffield Council on Bioethics. Artificial intelligence (AI) in healthcare and research. London: Nuffield Council on Bioethics. 2018; 8 p.
7. Gödel K. On Formally Undecidable Propositions of Principia Mathematica and Related Systems I (1931). New York: Dover Publications Inc. 1992; 79 p.
8. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, et al. AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines. 2018; (28): 689–707.
9. Siau K, Wang W. Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI. Journal of Database Management. 2020; 2 (37): 74–87.
10. Kozhevnikova M, Karpova S. V. Iskusstvennyj intellekt: sub#ekt i ob#ekt. Jethnograficheskoe obozrenie. 2020; (1): 80–94. Russian.
11. Anderson SL, Anderson M. AI and ethics. AI and Ethics. 2021; 1 (1): 27–31.
12. Panetta K. Trends Drive the Gartner Hype Cycle for Emerging Technologies [Internet]. 2020 [cited 2021 April 05]. Available from: <https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020/>