

DOI: <https://doi.org/10.17816/DD430334>

Концепция ответственного искусственного интеллекта — будущее искусственного интеллекта в медицине

Н.С. Германов

Российский национальный исследовательский медицинский университет имени Н.И. Пирогова, Москва, Российская Федерация

АННОТАЦИЯ

Активное развертывание систем искусственного интеллекта (ИИ) в медицине создаёт множество трудностей. В последние годы широко обсуждается концепция ответственного искусственного интеллекта (ОИИ), направленная на решение неизбежно возникающих этических, юридических и социальных проблем. Проведён анализ научной литературы по теме, рассмотрена возможность применения концепции ОИИ для преодоления существующих проблем ИИ в медицине. Исследования возможных применений ИИ в медицине показали, что современные алгоритмы не способны удовлетворить базовые непреходящие потребности общества — справедливость, прозрачность и надёжность. С целью решения данных этических проблем была предложена концепция ОИИ, основывающаяся на трёх принципах (ART) — объясняемость и прозрачность выводов, а также ответственность за деятельность ИИ. Дальнейшее развитие, без разработки и применения ART-концепции, делает опасным и невозможным применение ИИ в таких областях, как медицина, государственное управление и т.д. Требования к объясняемости и прозрачности выводов основываются на выявленных эпистемологических (ошибочные, непрозрачные, неполные выводы) и нормативных (конфиденциальность данных, дискриминация определённых групп) проблемах применения ИИ в цифровой медицине [2]. Эпистемологические ошибки, совершаемые ИИ, не ограничиваются упущениями, связанными с объёмами и репрезентативностью исходных анализируемых баз данных. К ним также относят известную проблему «чёрного ящика» — невозможность «заглянуть» в процесс формирования выводов ИИ при обработке входных данных. Наряду с эпистемологическими ошибками также неизбежно возникают нормативные проблемы — конфиденциальность пациентов и дискриминация некоторых социальных групп вследствие отказа части пациентов на предоставление их медицинских данных для тренировки алгоритмов и в составе анализируемых баз данных, что приведёт к недостаточно точным выводам ИИ в случаях определённых гендеров, рас, возрастов и т.д.

Важно отметить, что методология анализа данных ИИ зависит от кода программы, заданного программистом, эпистемологические и логические ошибки которого неизбежно проецируются на ИИ. Отсюда следует проблема определения ответственности в случае ошибочных выводов — её распределение между самой программой, разработчиком и исполнителем. Многочисленные профессиональные ассоциации проектируют этические стандарты для разработчиков и законодательную базу для регулирования ответственности между описанными звеньями. Однако очевидно, что в разработке и утверждении такого законодательства наибольшую роль должно играть государство.

Применение ИИ в медицине, несмотря на преимущества, сопровождается множеством этических, юридических и социальных проблем. Разработка ОИИ способна не только разрешить данные трудности, но и обеспечить дальнейшее активное и безопасное развертывание систем ИИ в цифровой медицине и здравоохранении.

Ключевые слова: искусственный интеллект; ответственный искусственный интеллект; цифровое здравоохранение; цифровая медицина; машинное обучение.

КАК ЦИТИРОВАТЬ

Германов Н.С. Концепция ответственного искусственного интеллекта — будущее искусственного интеллекта в медицине // *Digital Diagnostics*. 2023. Т. 4, № 1 Supplement. С. 27–29. DOI: <https://doi.org/10.17816/DD430334>

СПИСОК ЛИТЕРАТУРЫ

1. Dignum V. Responsibility and Artificial Intelligence. In: Dubber M.D., Pasquale F., Das S., editors. *The Oxford Handbook of Ethics of AI*. Oxford : Oxford University Press, 2020. doi: 10.1093/oxfordhb/9780190067397.013.12
2. Trocin C., Mikalef P., Papamitsiou Z., et al. Responsible AI for Digital Health: a Synthesis and a Research Agenda // *Inf Syst Front*. 2021. Режим доступа: https://www.researchgate.net/publication/352769689_Responsible_AI_for_Digital_Health_a_Synthesis_and_a_Research_Agenda/link/60d807df92851ca9448cf7c4/download. Дата обращения: 03.06.2023. doi: 10.1007/s10796-021-10146-4
3. Racine E., Boehlen W., Sample M. Healthcare uses of artificial intelligence: Challenges and opportunities for growth // *Healthcare Management Forum*. 2019. Vol. 32, N 5. P. 272–275. doi: 10.1177/0840470419843831
4. Zednik C. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence // *Philos. Technol*. 2021. Vol. 34. P. 265–288. doi: 10.1007/s13347-019-00382-7
5. Astromskė K., Peičius E., Astromskis P. Ethical and legal challenges of informed consent applying artificial intelligence in medical diagnostic consultations // *AI & Soc*. 2021. Vol. 36. P. 509–520. doi: 10.1007/s00146-020-01008-9
6. Burr C., Taddeo M., Floridi L. The Ethics of Digital Well-Being: A Thematic Review // *Sci Eng Ethics*. 2020. Vol. 26. P. 2313–2343. doi: 10.1007/s11948-020-00175-8
7. Gotterbarn D., Bruckman M., Flick C., Miller K., Wolf M.J. ACM Code of Ethics: A Guide for Positive Action // *Communications of the ACM*. 2018. Vol. 61, N 1. P. 121–128. doi: 10.1145/3173016

DOI: <https://doi.org/10.17816/DD430334>

The concept of responsible artificial intelligence as the future of artificial intelligence in medicine

Nikolai S. Germanov

N.I. Pirogov Russian National Research Medical University, Moscow, Russian Federation

ABSTRACT

Active deployment of artificial intelligence (AI) systems in medicine creates many challenges. Recently, the concept of responsible artificial intelligence (RAI) was widely discussed, which is aimed at solving the inevitable ethical, legal, and social problems. The scientific literature was analyzed and the possibility of applying the RAI concept to overcome the existing AI problems in medicine was considered. Studies of possible AI applications in medicine showed that current algorithms are unable to meet the basic enduring needs of society, particularly, fairness, transparency, and reliability. The RAI concept based on three principles — accountability for AI activities and responsibility and transparency of findings (ART) — was proposed to address ethical issues. Further evolution, without the development and application of the ART concept, turns dangerous and impossible the use of AI in such areas as medicine and public administration. The requirements for accountability and transparency of conclusions are based on the identified epistemological (erroneous, non-transparent, and incomplete conclusions) and regulatory (data confidentiality and discrimination of certain groups) problems of using AI in digital medicine [2]. Epistemological errors committed by AI are not limited to omissions related to the volume and representativeness of the original databases analyzed. In addition, these include the well-known “black box” problem, i.e. the inability to “look” into the process of forming AI outputs when processing input data. Along with epistemological errors, normative problems inevitably arise, including patient confidentiality and discrimination of some social groups due to the refusal of some patients to provide medical data for training algorithms and as part of the analyzed databases, which will lead to inaccurate AI conclusions in cases of certain gender, race, and age. Importantly, the methodology of the AI data analysis depends on the program code set by the programmer, whose epistemological and logical errors are projected onto the AI. Hence the problem of determining responsibility in the case of erroneous conclusions, i.e. its distribution between the program itself, the developer, and the executor. Numerous professional associations design ethical standards for developers and a statutory framework to regulate responsibility between the links described. However, the state must play the greatest role in the development and approval of such legislation. The use of AI in medicine, despite its advantages, is accompanied by many ethical, legal, and social challenges. The development of RAI has the potential both to solve these challenges and to further the active and secure deployment of AI systems in digital medicine and healthcare.

Received: 15.05.2023

Accepted: 05.06.2023

Published Online: 10.07.2023

Keywords: artificial intelligence; responsible artificial intelligence; digital healthcare; digital medicine; machine learning.

FOR CITATION

Germanov NS. The concept of responsible artificial intelligence as the future of artificial intelligence in medicine. *Digital Diagnostics*. 2023;4(1S):27–29. DOI: <https://doi.org/10.17816/DD430334>

REFERENCES

1. Dignum V. Responsibility and Artificial Intelligence. In: Dubber MD, Pasquale F, Das S, editors. *The Oxford Handbook of Ethics of AI*. Oxford: Oxford University Press; 2020. doi: 10.1093/oxfordhb/9780190067397.013.12
2. Trocin C, Mikalef P, Papamitsiou Z, et al. Responsible AI for Digital Health: a Synthesis and a Research Agenda. *Inf Syst Front*. 2021. Available from: https://www.researchgate.net/publication/352769689_Responsible_AI_for_Digital_Health_a_Synthesis_and_a_Research_Agenda/link/60d807df92851ca9448cf7c4/download. doi: 10.1007/s10796-021-10146-4
3. Racine E, Boehlen W, Sample M. Healthcare uses of artificial intelligence: Challenges and opportunities for growth. *Healthcare Management Forum*. 2019;32(5):272–275. doi: 10.1177/0840470419843831
4. Zednik C. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philos. Technol*. 2021;34:265–288. doi: 10.1007/s13347-019-00382-7
5. Astromskė K, Peičius E, Astromskis P. Ethical and legal challenges of informed consent applying artificial intelligence in medical diagnostic consultations. *AI & Soc*. 2021;36:509–520. doi: 10.1007/s00146-020-01008-9
6. Burr C, Taddeo M, Floridi L. The Ethics of Digital Well-Being: A Thematic Review. *Sci Eng Ethics*. 2020;26:2313–2343. doi: 10.1007/s11948-020-00175-8
7. Gotterbarn D., Bruckman M., Flick C, Miller K, Wolf MJ. ACM Code of Ethics: A Guide for Positive Action. *Communications of the ACM*. 2018;61(1):121–128. doi: 10.1145/3173016

ИНФОРМАЦИЯ ОБ АВТОРЕ

Германов Николай Станиславович;
ORCID: <https://orcid.org/0000-0003-1953-8794>;
e-mail: n.s.germanov@gmail.com

AUTHOR'S INFO

Nikolai S. Germanov;
ORCID: <https://orcid.org/0000-0003-1953-8794>;
e-mail: n.s.germanov@gmail.com