**Концепция ответственного искусственного интеллекта – будущее ИИ в медицине**

**АВТОРЫ**

Германов Н.С.1

1Федеральное государственное автономное образовательное учреждение высшего образования "Российский национальный исследовательский медицинский университет имени Н.И. Пирогова" Министерства здравоохранения Российской Федерации, г. Москва, ул. Островитянова, д. 1

***Обоснование*:**

Активное развертывание систем искусственного интеллекта (ИИ) в медицине создает множество трудностей перед обществом. В последние годы широко обсуждается концепция ответственного искусственного интеллекта (ОИИ), направленная на решение неизбежно возникающих этических, юридических и социальных проблем.

***Цель*:**

Рассмотреть возможность применения концепции ответственного искусственного интеллекта для преодоления существующих проблем ИИ в медицине.

***Методы*:**

Анализ научной литературы по теме.

***Результаты*:**

Исследования возможных применений ИИ в медицине показали, что современные алгоритмы не способны удовлетворить базовые непреходящие потребности общества – справедливость, прозрачность и надежность. С целью решения данных этических проблем была предложена концепция ОИИ, основывающаяся на трех принципах (ART) – объясняемость и прозрачность выводов, а также ответственность за деятельность ИИ. [1,2] 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 определенных групп) проблемах применения ИИ в цифровой медицине. [2]

Эпистемологические ошибки, совершаемые ИИ, не ограничиваются упущениями, связанными с объемами и репрезентативностью исходных анализируемых баз данных. К ним также относят известную проблему “черного ящика” – невозможность “заглянуть” в процесс формирования выводов ИИ при обработке входных данных. [3,4]

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 приведет к недостаточно точным выводам ИИ в случаях определенных гендеров, рас, возрастов и т.д. [5,6]

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 – ее распределение между самой программой, разработчиком и исполнителем. Многочисленные профессиональные ассоциации проектируют этические стандарты для разработчиков и законодательную базу для регулирования ответственности между описанными звеньями. [1,7] Однако очевидно, что в разработке и утверждении такого законодательства наибольшую роль должно играть государство.

***Заключение*:**

Применение ИИ в медицине, несмотря на свои 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.

**The concept of responsible artificial intelligence – the future of AI in medicine**

**AUTHORS**

Germanov N.S.1

**AFFILIATION**

1Pirogov Russian National Research Medical University, Moscow, Russia

**KEYWORDS**

Artificial intelligence, responsible artificial intelligence, digital healthcare, digital medicine, machine learning.

**Список литературы**

1. Dignum, Virginia, 'Responsibility and Artificial Intelligence', in Markus D. Dubber, Frank Pasquale, and Sunit Das (eds), The Oxford Handbook of Ethics of AI (2020; online edn, Oxford Academic, 9 July 2020), <https://doi.org/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). <https://doi.org/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](https://doi.org/10.1177/0840470419843831)
4. Zednik, C. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. Philos. Technol. 34, 265–288 (2021). <https://doi.org/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 36, 509–520 (2021). <https://doi.org/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 26, 2313–2343 (2020). <https://doi.org/10.1007/s11948-020-00175-8>
7. D. W. Gotterbarn et al., “ACM Code of Ethics: A Guide for Positive Action,” Communications of the ACM 61, no. 1 (2018): 121– 28.

-------------------------------------

Автор, ответственный за переписку – Германов Николай Станиславович, n.s.germanov@gmail.com

Германов Николай Станиславович

Germanov Nikolai Stanislavovich