От принципов к практике: внедрение ответственного управления ИИ в предиктивной и генеративной клинической ИИ-системах

Источник: Frontiers in Digital Health

Оригинал: https://www.frontiersin.org/articles/10.3389/fdgth.2026.1965076

генеративный ИИ | клинические испытания | предиктивная аналитика | регулирование | управление ИИ | этика ИИ

Ответственное управление искусственным интеллектом (ИИ) в здравоохранении все больше опирается на международные принципы, касающиеся прозрачности, подотчетности, безопасности, объяснимости и управления рисками на протяжении всего жизненного цикла. Тем не менее, сохраняется устойчивый разрыв в реализации между этими принципами на уровне организаций и конкретными мерами контроля, применяемыми при использовании ИИ-сервисов в клинической практике. В данной перспективной статье утверждается, что ответственный ИИ должен быть переведен в практическую плоскость на функциональных этапах, где формируются, объясняются и доводятся до сведения пользователей результаты работы клинического ИИ. Мы используем три функциональные области — «Возможности» (Capability), «Объяснимость» (Explainability) и «Коммуникация» (Communication) — в качестве организационной структуры для перевода существующих принципов управления в спецификации уровня приложений. Этот подход учитывает как предиктивный ИИ (PAI), так и генеративный ИИ (GAI), признавая при этом их различные профили рисков. Для предиктивного ИИ управление акцентирует внимание на определении целевых показателей и горизонта прогнозирования, калибровке, пороговых значениях, дрейфе данных, внешней валидации и нагрузке от оповещений. Для генеративного ИИ акцент делается на границах задач, отслеживаемости

источников, контроле галлюцинаций, управлении версиями, проверке человеком и ограничениях на обратную запись данных или автономные действия. Два примера приложений, связанных с диализом, иллюстрируют, как один и тот же словарь управления может по-разному воплощаться в различных модальностях ИИ. Операционализация ответственного ИИ таким образом не вводит новых принципов управления; скорее, она организует существующие требования в виде наблюдаемых, проверяемых и поддающихся аудиту мер контроля, встроенных в ИИ-сервисы и клинические рабочие процессы.

Перевод выполнен: 31.10.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.