От этики к доверию: стратегические ограничители для безопасного, защищённого и эффективного ИИ в здравоохранении

Источник: MedCity News

Оригинал: https://medcitynews.com/2026/08/from-ethics-to-trust-strategic-guardrails-for-safe-secure-effective-ai-in-healthcare/

безопасность пациентов | доверенный ИИ | здравоохранение | клинические workflows | разработка лекарств | регулирование | управление | этика ИИ

Внедрение ИИ в здравоохранении всё ещё отстаёт от других отраслей, а общественное доверие остаётся ограниченным, что усиливает потребность в стратегическом лидерстве, чётких ограничительных рамках и ответственном внедрении.

Внимание к ИИ растёт, как и скептицизм, и давление с целью сделать всё правильно, особенно понимая, что эти приложения могут влиять на жизни пациентов. По мере того как ИИ становится более интегрированным в разработку лекарств и клинические рабочие процессы, от лидеров в сфере здравоохранения и наук о жизни ожидается демонстрация этического управления и технической компетентности для укрепления доверия и защиты пациентов.

Именно поэтому этический ИИ и фреймворки доверенного ИИ должны быть базовыми требованиями для организаций здравоохранения и наук о жизни, стремящихся к инновациям с помощью ИИ. Эти различные, но взаимосвязанные фреймворки способствуют созданию защищаемой стратегии ИИ — такой, которая гарантирует, что все действия и решения,

связанные с ИИ, могут быть объяснены и обоснованы перед ключевыми заинтересованными сторонами. Вместе они позволяют командам ответить на два критических, проверенных давлением вопроса: какие риски связаны с этим приложением ИИ и способны ли мы их смягчить?

Различия между этическим ИИ и доверенным ИИ

Различие между «этическим ИИ» и «доверенным ИИ» имеет значение для фармацевтических руководителей, поставщиков медицинских услуг и регуляторов, которые влияют на то, как ИИ разрабатывается и внедряется. Хотя эти концепции могут пересекаться, полезно рассматривать их отдельно при установлении стратегии, ограничительных рамок и ожиданий относительно доказательств.

Этический ИИ и доверенный ИИ являются операционными фреймворками, которые формируют то, как организации управляют, внедряют и завоевывают доверие к решениям ИИ. Их взаимодействие и последовательность варьируются в зависимости от контекста и организационных приоритетов. На практике организации могут решать вопросы этического и доверенного ИИ одновременно, используя доверенный ИИ для определения того, что может быть разработано и оценено, в то время как этические соображения направляют надзор на протяжении всего жизненного цикла.

Вот главное различие: этический ИИ исследует, что **должно быть сделано**, отражая позицию руководства организации через ценности, нормы и справедливость. Этические фреймворки специализированы и балансируют потребности политики, технологий и бизнеса. Поскольку многие этические соображения включают нюансы и компромиссы, фреймворки должны быть адаптируемыми по мере появления новой информации или контекста.

Например, при использовании социальных детерминант здоровья (SDoH) для руководства распределением ресурсов с помощью ИИ для общественного здравоохранения, полагаемся ли мы на показатели на уровне сообщества, которые уважают контекст, или на сигналы на индивидуальном уровне, которые могут казаться навязчивыми или несправедливыми? При наборе в клинические исследования приемлемо ли оптимизировать ИИ для прогнозируемого завершения, если это снижает репрезентативность популяции, и где находится этически защищаемый баланс?

Доверенный ИИ исследует, что **может быть сделано**. Эти фреймворки фокусируются на технических требованиях (что нужно построить и разработать) и подчёркивают демонстрируемое соответствие, надёжность, безопасность, защищённость, прозрачность и объяснимость. Доверенный ИИ также учитывает обязательное законодательство и признанные фреймворки, такие как EU AI Act (законодательство) и NIST AI Risk Management Framework (руководство), которые консультируют организации по вопросам ответственной практики ИИ, формируя то, как принципы доверенного ИИ подтверждаются доказательствами.

Продолжая предыдущие примеры, если департамент общественного здравоохранения использует SDoH (возможно, на уровне сообщества), доверенная практика означает разработку с учётом безопасности пациентов и защиты данных, включая технические контроли и паттерны доступа, необходимые для соблюдения законов об ИИ и защите данных, соответствие соответствующим стандартам и создание проверяемых доказательств этих контролей. Для набора в исследования, балансируя разнообразие популяции с вероятностью завершения исследования в усилиях по набору, фокус будет включать точные прогнозы завершения across всех подпопуляций. Стратегия набора также должна оставаться клинически обоснованной и защищаемой перед регуляторами, такими как FDA (Food and Drug Administration, Управление по санитарному надзору за качеством пищевых продуктов и медикаментов США) или EMA (European Medicines Agency, Европейское агентство по лекарственным средствам).

Связная модель

Понимание различий между этическим и доверенным ИИ — это одно. Воплощение их в практику — другое. Вопрос становится таким: как нам операционализировать оба фреймворка одновременно?

Этический ИИ и доверенный ИИ должны переплетаться под связным стратегическим надзором, обеспечивая чёткие ограничительные рамки, роли и подотчётность, которые адаптируются по вариантам использования. Разработчики лекарств, исследующие использование ИИ для обнаружения нежелательных явлений, будут иметь другие риски и другие рабочие процессы, чем больница, использующая ИИ для клинической документации. Но обоим нужны структуры подотчётности, которые доказывают, что использование ИИ является этическим, доверенным, ответственным и защищаемым.

Воплощение этики и доверия в работу

Создание связного надзора и чёткого направления требует преднамеренного сотрудничества. Регулирование и общественные ожидания эволюционируют across отраслей и регионов. Лидеры могут снизить риски прогресса, устанавливая стратегические ограничительные рамки рано, измеряя результаты безопасности и защищённости, и обновляя доказательства по мере созревания стандартов.

Чтобы соответствовать повышенному scrutiny и продемонстрировать стратегический надзор, прозрачность и подотчётность, организации могут разработать собственные фреймворки, которые балансируют операционную эффективность с безопасностью пациентов, защищённостью и этическими ограничительными рамками. Вот некоторые соображения для начала:

- **Установите ИИ как стратегический приоритет**: Намерение, влияние и принципы технического дизайна должны быть определены с самого начала, чтобы цель, безопасность пациентов и защищённость были встроены, а не добавлены позже. Установление этически ведомой стратегической позиции (почему) наряду со стандартами доверенной разработки (что) даёт командам чёткие ограничительные рамки и направление с первого дня. Междисциплинарное лидерство (включая клинических специалистов, продукт, науку об ИИ, безопасность, конфиденциальность, юридические вопросы) совместно владеет этой позицией и развивает её по мере созревания доказательств и регулирования. Работа таким образом снижает сюрпризы на поздних стадиях и согласовывает пилотные проекты и разработку с культурой и ожиданиями здравоохранения и наук о жизни. Это помогает снизить проблемы на более поздних этапах, поддерживает внедрение и снижает операционный риск.

- **Формализуйте ограничительные рамки и направление**: Стратегический надзор поддерживает инновации, создавая ясность о том, как ИИ рассматривается и применяется. В здравоохранении и науках о жизни формальный совет по надзору за этикой и доверием, который объединяет старших руководителей и технических экспертов, может обеспечить последовательный способ оценки деятельности, выявления пробелов в доказательствах и решения вопросов о подотчётности. Эта структура помогает командам рано выявлять варианты использования с высоким риском или неоднозначные и применять соответствующий уровень scrutiny с

безопасностью пациентов и защищённостью как первичными принципами. Когда надзор работает параллельно с рабочими потоками внедрения, организации решают регуляторные обязанности, риски и этические вопросы как часть построения, поэтому ответственное инноватирование движется быстрее и с уверенностью.

- **Используйте многоуровневый подход к риску**: Организация здравоохранения или наук о жизни, поддерживаемая экспертами по конфиденциальности и соответствию в совете по рассмотрению с исполнительным надзором, должна оценивать варианты использования ИИ через многоуровневый подход к риску — тот, который учитывает тип ИИ, а также его позицию в жизненном цикле продукта и близость к жизням пациентов. Этот адаптивный подход масштабирует оценку риска ИИ и достоверности на основе потребности (где ограничительные рамки уже были внедрены для многих приложений). Например, интеллектуальный анализ литературы и моделирование сценариев являются приложениями ИИ с более низким риском, используемыми в исследованиях на ранней стадии, что также считается более низким риском. Тогда как приложения с более высоким риском, такие как большие языковые модели, используемые в поддержке клинических решений, требуют более тщательной оценки из-за их сложности и близости к уходу за пациентами. Безопасность пациентов является императивом и линзой, через которую принимаются технические и бизнес-компромиссы.

- **Создайте культурную готовность**: Это является основной частью эффективного внедрения ИИ, потому что командам нужно понимать, как технология вписывается в их работу и как принимаются решения. Программы ИИ набирают обороты, когда лидеры объясняют, как приложения поддерживают этические обязательства, результаты безопасности пациентов, стандарты защищённости и бизнес-цели. Лидерам также нужно будет адресовать вопросы об изменениях рабочих процессов, регуляторных ожиданиях или воспринимаемом влиянии на работу с ясностью. Предоставление структурированной поддержки управления изменениями помогает людям увидеть, как ИИ может положительно способствовать операционным целям и результатам для пациентов, что создаёт уверенность и позволяет внедрение across организации.

Потребность в тщательном внедрении постоянна, потому что ИИ в здравоохранении и науках о жизни неизбежно влияет на пациентов и их уход. Со стратегическим надзором, чёткими ограничительными рамками, мерами безопасности пациентов и безопасной разработкой организации могут применять ИИ ответственно, завоевывая доверие клиницистов, регуляторов и, что важнее всего, пациентов.

Luk Arbuckle является глобальным лидером практики ИИ и главным методологом в IQVIA Applied AI Science, двигая инновации в данных и ИИ с фокусом на этику, конфиденциальность и регуляторное соответствие. Признанный лидер мысли, Luk консультирует глобальные органы власти и широко публикуется по вопросам управления ИИ и защиты данных.
