

Реальная проблема всех нас — не согласие, а конвертация данных

Источник: STAT

Дата публикации: 2026-09-15

Оригинал: <https://medcitynews.com/2026/10/the-real-gap-in-all-of-us-isnt-consent-its-conversion/>

интероперабельность

исследовательские базы данных

обработка естественного языка

структурирование данных

электронные медкарты

Недавно STAT сообщило, что, несмотря на то, что программа исследований «All of Us» Национального института здоровья (NIH) убедила 98% своих почти 750 000 участников своими электронными медицинскими картами (ЭМК), более 300 000 этих участников вообще не имеют данных ЭМК в базе данных программы. Последняя публикация данных программы пытается закрыть эту брешь, используя те же сети обмена клиническими данными, которые больницы уже используют для перемещения записей между системами здравоохранения.

Это разумное решение, и стоит остановиться на том, почему оно вообще необходимо, поскольку объяснение говорит о чем-то важном о том, где на самом деле находится американская инфраструктура медицинских данных в 2026 году, и это не там, где большинство людей предполагает.

В течение десятилетия рабочая теория интероперабельности (совместимости систем) сводилась по сути к проблеме сантехники: подключи трубы, и данные потекут. Эта теория была в значительной степени верной, и сантехника, действительно, была построена. Правила информационной

блокировки Закона о науке и технологиях XXI века (21st Century Cures Act) закрепили право пациента на собственные данные. Сертифицированные поставщики ЭМК были обязаны предоставлять стандартизированные API в формате FHIR (Fast Healthcare Interoperability Resources). Национальные сети теперь позволяют больным запрашивать записи друг друга с легкостью, близкой к телефонному справочнику. Люди, которые выполняли эту работу в федеральных офисах по политике ИТ в здравоохранении, в системах здравоохранения и среди поставщиков ЭМК, совершили нечто по-настоящему сложное, преодолев реальное институциональное сопротивление, и это отражается в таких цифрах, как 98% уровень согласия в программе «All of Us». Пациенты готовы. Провода в основном подключены.

Так почему же 300 000 записей до сих пор не отображаются?

Потому что получение записи и наличие пригодных для использования данных из нее — это две разные проблемы, и отрасль решила первую, не завершив вторую. Больница теперь, в принципе, может передать полную карту пациента в базу данных исследований за секунды. Однако то, что приходит, редко бывает чистым структурированным набором данных. Это смесь: некоторые дискретные ресурсы FHIR, некоторые отсканированные направления, некоторые факсимильные формы предварительного авторизования, некоторые экспорты в формате PDF с записями о визитах, поля в которых различаются от больницы к больнице, от ЭМК к ЭМК, от десятилетия к десятилетию. Карта пациента может охватывать несколько разных систем, несколько разных форматов документов и годы лечения, проведенные до того, как кто-либо подумал сделать записи машиночитаемыми. Труба работает. Содержимое, текущее по ней, часто не приходит в форме, которую может использовать любая база данных для исследований.

Это та часть интероперабельности, о которой редко говорят в политических дискуссиях, потому что это не проблема политики. Это проблема инженерии данных; и, к счастью, это проблема, которая решается гораздо ближе, чем обычно предполагают большинство руководителей систем здравоохранения, исследователей или даже технологических специалистов.

Инструменты для преобразования по-настоящему беспорядочных, несогласованных документов реального мира в структурированные, проверенные данные, готовые для исследований, существенно созрели за последние два года. То, что делает это возможным, — это не прорыв одной модели. Это сочетание вероятностного извлечения (системы ИИ, которые могут прочитать отсканированный, рукописный или несогласованно

отформатированный документ примерно так, как это сделал бы обученный человек) и детерминированных слоев проверки, которые проверяют каждое извлеченное поле на соответствие определенным правилам, прежде чем оно будет принято в набор данных. Это сочетание важно, потому что одного вероятностного ИИ недостаточно для обеспечения надежности на порогах точности, требуемых для исследований и клинического использования, а системы, основанные только на правилах, не могут справиться с огромным разнообразием форматов реальных записей. Вместе они могут. И разрыв между двумя этими подходами, который раньше измерялся годами кастомной разработки для каждой системы здравоохранения, сейчас сокращается за недели.

Почему фармацевтическим компаниям следует обращаться к социальным сетям за лучшими знаниями о пациентах

Лица, принимающие решения в сфере здравоохранения, особенно в фармацевтических компаниях, должны придавать первостепенное значение мнению пациентов. Данные, собранные с помощью мониторинга социальных сетей, являются значительным шагом в правильном направлении.

Это важно по трем причинам, которые напрямую относятся к тому, что на самом деле нужно программе «All of Us» и любой другой исследовательской программе, системе здравоохранения или плательщику, сталкивающемуся с той же проблемой 300 000 записей.

Это быстрее, чем сроки, на которые рассчитаны большинство учреждений. Превращение отложенных в долгий ящик неструктурированных записей в структурированные, пригодные для использования данные больше не требует многолетнего проекта интеграции или замены устаревших систем. Это может происходить параллельно с уже существующими системами, потребляя любой формат, в котором поступает запись (факс, скан, PDF или структурированный экспорт), не требуя от системы здравоохранения изменения того, как или где она что-либо хранит.

Это более точно, чем ручная альтернатива, а не менее. Инстинкт в исследовательских и регуляторных средах заключается в том, чтобы рассматривать автоматизацию как компромисс с точностью, считая ручную абстракцию карт безопасным выбором по умолчанию. На практике детерминированные слои проверки могут пометить каждое поле, выходящее за пределы ожидаемых границ, для ручного обзора, создавая задокументированный аудиторский след, показывающий, что именно было

извлечено, из какого источника и с каким уровнем уверенности, демонстрируя более строгую доказательную базу, чем большинство современных процессов ручной абстракции.

И это совместимо со стандартами безопасности и происхождения данных, которым должна соответствовать такая программа, как «All of Us». Ничто из этого не требует, чтобы записи покидали контролируемые среды, подвергая данные потребительским инструментам ИИ, или ослабляли позиции соответствия, такие как SOC 2 и HIPAA, которым уже подчиняются регулируемые исследовательские программы. Слой проверки, по сути, является тем, что дает аудиторам и институциональным наблюдательным советам отслеживаемость, необходимую им для доверия автоматическому извлечению данных в первую очередь.

Ничто из этого не является критикой подхода «All of Us». Использование существующих сетей обмена данными для получения большего количества записей — это правильный шаг, и он закроет часть бреши. Но в конечном итоге он упрется в ту же стену, с которой сталкивается каждое крупномасштабное усилие по извлечению данных из ЭМК: значительная часть того, что возвращается, не будет чистыми структурированными данными. Это будет беспорядочный остаток, отсканированный документ, старый экспорт или запись из системы, которая никогда полностью не приняла FHIR, что одно лишь извлечение исправить не может.

Честный урок из бреши в 300 000 записей заключается не в том, что интероперабельность потерпела неудачу. Он заключается в том, что интероперабельность, как ее в основном определила отрасль (может ли запись переместиться из системы А в систему В), добилась значительного успеха, а оставшаяся работа — это другая, более решаемая проблема, чем та, над которой отрасль работала последние десять лет. Превращение того, что приходит, в то, что исследователи могут фактически использовать, больше не является многолетним предприятием, каким оно считалось пять лет назад. Это ближе к решенной инженерной проблеме, чем осознает большинство людей, создающих инфраструктуру медицинских данных. Исследовательские программы, которые будут относиться к этому как к решенной проблеме, закроют такие бреши гораздо быстрее, чем это сделала работа по разработке политики за последнее десятилетие.

Фото: Bigstock

Адам Коэн — соучредитель компании Morph Services, Inc., технологической компании в сфере здравоохранения, которая преобразует неструктурированные клинические документы в надежные структурированные данные для систем здравоохранения, плательщиков и других регулируемых организаций. Имея более 15 лет опыта, охватывающего операции в здравоохранении, интероперабельность, обработку данных с использованием ИИ, аккредитацию и корпоративное программное обеспечение, он помог организациям модернизировать сложные рабочие процессы, сохраняя точность, управление и соответствие требованиям, необходимые в строго регулируемых средах. До соучреждения Morph Коэн занимал руководящие должности в технологических компаниях здравоохранения, специализирующихся на телемедицине, аккредитации и цифровой трансформации, поддерживая многие из крупнейших систем здравоохранения и медицинских организаций страны. Он является частым автором и спикером по вопросам интероперабельности в здравоохранении, ИИ и практических проблем преобразования фрагментированной клинической информации в данные, пригодные для использования, которые улучшают уход, исследования и операционную эффективность.

Эта статья опубликована в рамках программы MedCity Influencers. Любой может опубликовать свое мнение о бизнесе и инновациях в здравоохранении на MedCity News через программу MedCity Influencers. Нажмите [здесь](#), чтобы узнать, как это сделать.

Перевод выполнен: 03.10.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.