

# Перспективы применения больших данных в российском здравоохранении

Александр Гусев



*Цифровизация, большие данные открывают новую страницу во всех сферах жизни общества. Пожалуй, одной из наиболее активно развивающихся в этом направлении отраслей является медицина и здравоохранение. Как это будет работать и как работает сейчас, рассказывает директор по развитию проекта Webiomed Александр Гусев.*

**Александр Гусев, кандидат технических наук, директор по развитию проекта Webiomed, эксперт по искусственному интеллекту Центрального научно-исследовательского института организации и информатизации здравоохранения Минздрава России, старший научный сотрудник Научно-практического клинического центра диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, ответственный редактор журнала «Врач и информационные технологии»**

## — Что такое большие данные применительно к здравоохранению?

— Как такового устоявшегося определения термина «большие данные» пока нет. Вообще термин появился в 2008 году и пришел из телекома на фоне бурного развития отрасли. «Большими» предлагали называть данные, ежедневный объем которых составлял больше 150 гигабайтов. Именно такие гигантские объемы информации тогда появились в телекоммуникациях — сотовая связь, интернет. Постепенно термин стал применяться

практически везде, в том числе и в здравоохранении. Сейчас большими данными стали называть такой способ их сбора, который хранит полностью картину какого-то объекта. Например, если мы в одну папку запишем всю медицинскую информацию пациентов региона или даже крупного медицинского центра, включая генетические данные, всю родословную, все факторы окружающей среды, это тоже будут гигабайты информации. И даже один человек — это уже тоже «большие данные». Не только и не столько по размеру, сколько по всеохватности. Или большие данные клиники, когда в одной информационной системе есть информация обо всех протекающих рабочих процессах: о ее пациентах, врачах, финансовых потоках, лекарственных ресурсах, оборудовании и его загруженности и т. д. Так что термин не устоялся, хотя слово употребляется очень широко. В любом случае, как мы считаем, большие данные — это колоссальное количество либо записей, либо объема информации.

Тем не менее все эксперты сходятся во мнении, что у этого явления — «большие данные» — есть явные черты. Для работы с большими данными необходимы



специализированные средства. Обычная настольная база данных не выдержит нагрузки и не сможет хранить и обрабатывать запросы на доступ к этим данным. Поэтому, как правило, речь идет о специализированных системах управления данными. Обычные программные продукты, например медицинская информационная система, могут храниться в реляционной СУБД (система управления базами данных), но если мы собираем данные, например, по целому региону, где все централизовано, то требуется уже другой специализированный инструментарий и, следовательно, специализированные методики работы с большими данными, потому что при таком колоссальном объеме невозможно проанализировать каждую строчку или, например, подсчитать в онлайн-режиме неподготовленные запросы. Такие большие данные, как правило, готовят к аналитической обработке. Это радикально отличает их от обычных данных в обычных информационных системах.

Применительно к здравоохранению инструментарий управления, организации процессов смещается к автоматической обработке объективных данных, которые мы собираем в реальном мире: в медицинских системах клиник, лабораторных системах, социальных сетях — это истинно большие данные, терабайты информации. Анализируя их в режиме реального времени, мы действительно можем увидеть реальную картину мира и принимать оперативные и иногда стратегические решения по отрасли на основе такого инструмента.

### — Каковы перспективы больших данных в медицине и здравоохранении, конкретные возможности применения?

— Сегодня большие данные плюс технологии искусственного интеллекта как способ извлекать из этих данных ценность очень напоминают ситуацию с появлением электричества. Всем известно, что в результате его открытия произошла индустриальная революция. Но если внимательно изучить то время, мы увидим, что сценариев применения, конкретной пользы электричества было очень мало: лампочка, освещение улиц, потом связь... Технология появилась, но ее применение, а затем и тотальная замена старого, устоявшегося индустриального подхода потребовали времени. Предприниматели, промышленники

с каждым годом создавали на основе все той же технологии все новую и новую ценность, но я уверен, что изобретатели электричества даже не представляли себе электромобилей и всех бытовых приборов, которыми мы сейчас пользуемся не задумываясь. Большие данные и искусственный интеллект в здравоохранении — точно такая же фундаментальная технология. Мы понимаем, что границ применения и потенциальных эффектов мы сейчас просто не видим. Пока мы довольно существенно ограничены и находимся на самых начальных этапах освоения больших данных в практическом здравоохранении. Нам доступна только верхушка айсберга, но и она уже впечатляет. Поэтому сомнений нет, что нужно накапливать и накапливать эти большие данные, даже если мы еще до конца не научились извлекать из них выгоду. Чем больше будет этих данных, тем большей ценностью

будут обладать технологии искусственного интеллекта. К тому же, и это самое интересное, мы можем создавать новую ценность, новые цифровые проекты.

### — Что вкладывается в понятие «управление большими данными»?

— Чтобы не было подмены понятий, необходимо четко разделить термины «управление большими данными» и «управление на основе данных». Что касается первого, это техническая наука, сфера информационных техно-

логий, которая изучает и разрабатывает в основном прикладные решения: конкретные программные продукты, математические алгоритмы, аппаратное обеспечение, для того чтобы мы могли обрабатывать и накапливать данные. Сегодня проблема больших данных состоит в том, что они генерятся с такой скоростью, что компьютеры или программное обеспечение, не говоря уже о людях, не всегда способны их обработать. Даже сверхпроизводительные процессоры и высоконагруженные приложения уже не всегда справляются.

Управление на основе данных — это наука об организации процессов по принятию управленческих решений на основе анализа данных, так называемый data-based management. Корпоративный, зрелый бизнес давно живет по этим правилам. Они означают, что решения по процессам не принимаются, если эти процессы не измерены, если нет объективной, собранной без участия человека и его интерпретаций информации. Решения принимаются только по тем процессам, которые

## БОЛЬШИЕ ДАННЫЕ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ЗДРАВООХРАНЕНИИ — ФУНДАМЕНТАЛЬНАЯ ТЕХНОЛОГИЯ. ГРАНИЦ ЕЕ ПРИМЕНЕНИЯ И ПОТЕНЦИАЛЬНЫХ ЭФФЕКТОВ МЫ СЕЙЧАС ПРОСТО НЕ ВИДИМ

автоматизированы и работают на цифровых платформах. Только такие процессы можно считать управляемыми. Подобные тезисы кажутся иногда наивными или даже оторванными от реальности. Но именно когда эту теорию изучаешь, когда она внедряется в организациях, приходит понимание, как работают эти аналитические порты, как формируются прогнозные модели, строятся тренды, создается красивая визуализация, а вы сосредоточены не на том, чтобы копаться в отчете, а на том, чтобы искать причину того или иного явления и пути решения. Методики управленца, построенные на анализе накопленных данных, позволяют выиграть в конкурентной борьбе, выжить в условиях ограниченных ресурсов по времени, финансам. Руководителю не надо верить кому-то на слово, он работает с информационными системами.

### — Как реализуется это в системах здравоохранения?

— В основном эти технологии научились применять для решения вполне конкретных, прикладных задач. Несколько примеров. Накопление больших данных позволило разрабатывать системы автоматического контроля за лечебно-диагностическим процессом, системы поддержки принятия врачебных решений, которые, анализируя большие данные, могут подсказывать врачу об ошибке или рекомендовать те или иные исследования, терапию. У нас миллионы пациентов, и системы учатся находить среди них подобных, выявлять закономерности, упрощая врачу процесс диагностики и лечения. Во-вторых, большие данные очень хорошо используются в научных исследованиях: новые данные о влиянии факторов риска, симптомах заболеваний, каких-то нетипичных течениях. Все эти крупницы знаний добавляются в медицинскую науку и шлифуются. В-третьих, много интересных кейсов в организации и управлении. Например, показательный кейс одной американской технологической компании, которая к определенной категории пациентов направляла медсестер для проведения обследования прямо на дому. По результатам пациенту назначалась терапия, амбулаторное лечение. Цель была — определить: болен / не болен, лечить / не лечить. Другая компания делала то же самое, но собрала шире данные, разработала алгоритмы анализа этих данных и в результате смогла стратифицировать этих пациентов

прямо дома по группам риска и выявить тех, у кого могут быть серьезные проблемы со здоровьем. Медицинская помощь уже оказывалась в зависимости от группы риска. Таким образом компания смогла на 15 % сократить госпитализации своих пациентов. Если учесть, что рынок здравоохранения США составляет несколько триллионов долларов, можно себе представить, что сокращение на 15 % госпитальных затрат — это колоссальный экономический эффект, выгода для работодателей, для пациентов, для здравоохранения.

### — Получается, что выгоднее оказалось буквально ко всем после определенного возраста отправлять на дом средний медперсонал, чтобы обследовать людей, выявить эту группу и ее контролировать?

— Совершенно верно. Рутинный скрининг в сочетании с использованием больших данных, а следовательно, прогнозной аналитики, машинного обучения — это очень серьезный способ оптимизации работы системы здравоохранения.

Управление на основе данных — это будущее организации здравоохранения, возможность прогнозировать ситуацию, принимать управленческие решения, перераспределять ресурс не после того, как случилась катастрофа, а заранее. И это чуть ли не единственный путь, потому что заболеваемость растет, траты на здравоохранение повышаются, ведь чем дольше живет

пациент, тем больше у него хронических заболеваний, тем больше он нуждается в медицинской помощи, а значит, ресурсов будет не хватать все больше, и это глобальная проблема. Сбор данных и принятие управленческих решений на их основе — это способ обеспечить эффективность существующей системе. Альтернативный путь — резкое увеличение затрат на здравоохранение. Но если львиная доля ресурсов пока еще идет на предотвратимую медицинскую помощь, к чему наращивать неэффективные траты? Ведь они могут увеличиваться только за счет других секторов экономики. Я убежден, что выход — это управление на основе цифровизации и больших данных.

### — А что в этом направлении уже реализуется в нашей стране?

— У нас есть повод для национальной гордости. Например, около 31 % стран мира, согласно аналитике ВОЗ,

## РУТИННЫЙ СКРИНИНГ В СОЧЕТАНИИ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ДАННЫХ, ПРОГНОЗНОЙ АНАЛИТИКИ, МАШИННОГО ОБУЧЕНИЯ – ЭТО СЕРЬЕЗНЫЙ СПОСОБ ОПТИМИЗАЦИИ РАБОТЫ СИСТЕМЫ ЗДРАВООХРАНЕНИЯ



представляет государственное стимулирование для внедрения электронных медицинских карт. И мы входим в это продвинутое меньшинство. У нас с 2011 года развивается проект Единой государственной информационной системы в сфере здравоохранения (ЕГИСЗ). Это колоссальное федеральное инвестирование, создана инфраструктура, куплено огромное количество компьютеров, созданы локальные вычислительные сети, территория у нас огромная, и потребности соответствующие, вложены огромные ресурсы в защиту информации, медицинское программное обеспечение. Например, по такому показателю, как доля медицинских организаций, внедривших систему ведения электронных медицинских карт, если не ошибаюсь, по итогам 2021 года мы вышли на уровень 90 % по всей стране. Это очень хороший показатель. В США наилучший в мире показатель — 98 %. 100 % нет нигде. А множество стран вообще только в самом начале этого пути. Но несмотря на такие успехи, у нас пока маленький рынок цифрового здравоохранения — 0,1–0,2 % от глобального, то есть потенциал роста огромный. Глобально пока лидируют Китай, Великобритания, США. В России — традиционно хорошие идеи, успешно создаются системы, но пока слабый предпринимательский задел.

### — Немного футурологии: как будет выглядеть система здравоохранения, управляемая на основе больших данных?

— Банальная истина: любое будущее невозможно предсказать. Но анализируя спокойно и трезво, через экспертную дискуссию, мы можем вырабатывать сценарии, выбирать из них, как нам кажется, наиболее реалистичный и двигаться исходя из него. Если внимательно посмотреть на структуру инвестиций в цифровое здравоохранение, на то, какие продукты и для кого создаются в последние два-три года, мы увидим тихую революцию, которая происходит у нас за спиной каждый день. Еще пять лет назад цифровое здравоохранение главным своим потребителем, плательщиком и пользователем видел участника здравоохранения. Создавались продукты для клиник, которые покупал главный врач или директор, для фарминдустрии, для применения в стационарах, поликлиниках. Сейчас львиная доля инвестиций направляется в продукты для пациентов. Компании, работающие в сфере цифрового здравоохранения, классическую

клинику, медицинскую организацию, врача и руководителя уже воспринимают не только как своего заказчика, но и как своего потенциального конкурента. Примерно через 5–10 лет мы окажемся в ситуации, когда цифровое здравоохранение станет «перехватывать» пациента у обычной медицинской организации, работающей в офлайн-режиме, и будет стремиться к тому, чтобы он туда не попал, не заболел или, по крайней мере, у него не развилось осложнение. Вот это один из самых интересных сценариев будущего, с моей точки зрения. Превентивное здравоохранение станет виртуальным и нацеленным на то, чтобы не позволить пациенту попасть под нож хирурга или просто на койку в стационар.

### — Речь идет о развитии инструмента, аналогичного системе поддержки принятия врачебных решений на уровне профилактики, чтобы пациент мог самостоятельно контролировать здоровье или процесс лечения? Но кто-то же должен проанализировать его личные данные, симптомы, показатели?

— Именно. Мы накапливаем большие данные, с их помощью мы знаем уже практически все о нашем пациенте, о течении его заболевания, можем прогнозировать, как оно будет развиваться, к чему приведет. Мы даже знаем порой, какие лекарства ему назначали и помогли они ему

или нет. Эти бесконечные знания спрятаны в цифре, и их надо оттуда извлечь. Как мы можем это сделать? Алгоритмы «интеллектуального помощника» сейчас активно обсуждаются. Концепция цифрового ассистента — это новый взгляд на систему поддержки принятия врачебных решений. На самом деле, это чисто российское название — «врачебных решений». На Западе устоявшийся термин — «система поддержки клинических решений» (clinical decision support system). Эти системы применяются в том числе средним медицинским персоналом, медсестрами, как основными и наиболее многочисленными участниками лечебно-диагностического процесса. Технологические компании сейчас говорят: мы можем вытащить из больших данных знания, чтобы создать виртуального ассистента, и почему он должен включаться в процесс только тогда, когда пациент свалился и оказался на койке в больнице, пусть он помогает ему дома, как только

## ЕСЛИ ВНИМАТЕЛЬНО ПОСМОТРЕТЬ НА СТРУКТУРУ ИНВЕСТИЦИЙ В ЦИФРОВОЕ ЗДРАВООХРАНЕНИЕ, НА ПРОДУКТЫ, КОТОРЫЕ СОЗДАЮТСЯ В ПОСЛЕДНИЕ ДВА-ТРИ ГОДА, МЫ УВИДИМ ТИХУЮ РЕВОЛЮЦИЮ

появился первый симптом и пациент еще даже не дошел до врача. Поэтому сейчас очень активно развиваются соответствующие мобильные приложения. Есть уже множество таких программ: вы сами можете сфотографировать родинку и получить с высокой долей точности ответ, есть ли онкологический риск; вбить симптомы, и алгоритм распознает, какое у вас, возможно, заболевание, скажет, куда обратиться и есть ли в этом необходимость. Цифровые ассистенты, обученные на больших данных, — это будущее.

### — Что требует развитие всех этих технологий от медицинского персонала? Расхожие опасения: что же будет делать врач?

— Это острый вопрос. Но основная задача — не создать совершенного виртуального врача, об этом даже не идет речь, а дать что-то пациенту. Технология — не конкурент врачу. Человечество уже прошло несколько технологических революций, и человек не пропал! Созданные им инструменты и машины не заменили его. Происходило исчезновение того человека, который хотел жить с каменным топором в эпоху электричества. Искусственный интеллект заменит только такого врача, который не применяет искусственный интеллект. Потому что это усиление его профессиональных возможностей, а не замена. Мы пришли не уничтожить, а повысить эффективность.

### — Как это влияет на медицинское образование?

— Для того чтобы наше образование не отставало от технологического уровня и производило кадры с актуальными компетенциями, цифровое здравоохранение нужно погружать в образовательный процесс, сотрудничать с такими авторами, как Татьяна Васильевна Зарубина и другие, общаться со студентами, находить среди них поклонников, интересантов. Тогда завтра врачи не будут относиться к «цифре» как к чему-то искусственно привнесенному технологической компанией. Сегодня для них абсолютная норма — ЭКГ и КТ, а завтра к ним добавится цифровой помощник.

### — Но ведь современные студенты уже естественным образом «прокачаны» в сфере цифровых технологий.

— Да, на мой взгляд, ничего дополнительного, искусственного вносить в жизнь не надо. Наоборот, роль академической науки состоит в том, что ей нужно это понять, принять и включать в образовательные программы, увеличивать часы медицинской кибернетики, давать практику по различным цифровым медицинским продуктам — естественным образом отвечать на запрос молодежи.

### — К вопросу о медицинской кибернетике. Она преподается сейчас только на специальных кафедрах. Имеет ли смысл сделать ее частью общего медицинского образования?

— Есть. Более того, в России несколько центров уже работают над этим. Например, Российский национальный исследовательский медицинский университет имени Н. И. Пирогова. Там сильная кафедра медицинской кибернетики, где готовят врачей-кибернетиков — не тех, кто будет пользоваться, а тех, кто будет правильно разрабатывать будущие продукты. Для врачей общей практики и других специальностей тоже есть базовые программы по компьютерной грамотности. Есть пласт профессионального сообщества в разных научных школах, разных научных центрах, сосредоточенный на том, чтобы обогащать образовательные программы цифровыми дисциплинами и для студентов, и для специалистов, и в последипломном образовании, и при переподготовке преподавателей.

### — Что вы можете сказать о Москве, о ее перспективах в работе с большими данными?

— У нас в стране сложились два концептуально разных подхода к цифровизации здравоохранения. Есть так называемый децентрализованный подход. Он подразумевает, что каждая клиника уникальна, только главный врач может принять решение, какой программный продукт ему выбрать, максимально заточивая его под свои процессы. Потом мы соединяем все эти клиники через определенные сервисы, чтобы получить общий банк данных этих клиник. По такому пути пошел Санкт-Петербург, сегодня это самый крупный децентрализованный проект в России. Его полный антагонист — централизованный подход, реализованный в Москве: у нас единый программный продукт, который внедряется во всей очень крупной сети медицинских организаций. За счет этого продукт развивается быстрее, мы сразу начинаем собирать данные централизованно, а не перегоняем их из системы в систему, форматируя и теряя иногда по пути. Москва, по моим ощущениям, — это один из крупнейших проектов в мире и по ведению электронного здравоохранения, электронной медицинской карты, электронной радиологии. ЕМИАС и ЕРИС в ее составе — это колоссальный производитель тех самых больших данных. И я верю, что в Москве зрелый менеджмент научится извлекать из этой «бигдаты» нужные данные и успешно решать конкретные вопросы эффективности. ММ