

Первичные данные ЭМК как «топливо» для систем поддержки принятия управленческих решений

Гусев Александр,

к.т.н., директор по развитию бизнеса, Webiomed

Мировое здравоохранение оцифровывается

2

Некоторые ключевые показатели

95%

МО ВНЕДРИЛИ ЭМК

В среднем среди развитых стран мира. Мы в России приближаемся к этому показателю

2314

ЭКЗАБАЙТ ДАННЫХ

Было произведено в сфере здравоохранения в 2020 году

48%

УВЕЛИЧЕНИЕ ДАННЫХ

Каждый год объем генерируемых здравоохранением данных растет на 48%

25

ТЫС. ПЕТАБАЙТ ДАННЫХ

Производит в среднем автоматизированная на данном этапе развития МО

1 трлн

УСТРОЙСТВ К 2035 г.

Будут собирать различную медицинскую информацию.

Направления сбора данных

- ✓ Электронные медицинские карты
- ✓ Результаты лабораторной диагностики
- ✓ Медицинские изображения
- ✓ Генетические данные
- ✓ Прием и движение лекарств
- ✓ Данные персональных медицинских устройств
- ✓ Данные мобильных приложений пациентов
- ✓ Данные социальных сетей
- ✓ Данные окружающей среды

Информатизация здравоохранения РФ

Рейтинг развития цифровой зрелости субъектов РФ
в сфере здравоохранения по итогам 2020 года

145 млрд

Гос. затраты
2011-2021 гг.

76%

Мед. организаций
внедрили МИС



● (90-100)% – 10 субъектов РФ

● (82-90)% – 17 субъектов РФ

● (65-82)% - 28 субъектов РФ

● (0-65)% - 29 субъектов РФ



МИНИСТЕРСТВО
ЗДРАВООХРАНЕНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Доклад «Реализация мероприятий по созданию цифрового контура в здравоохранении, направленных на повышение эффективности оказания медицинской помощи и снижению смертности от хронических и инфекционных заболеваний» зам. министра здравоохранения РФ П.С. Пугачева, 21.06.2021

Медицинская аналитика – крупнейший рынок

4

16

МЛРД. ДОЛЛ. США

Составил размер глобального рынка медицинской аналитики в 2019 г.

84

МЛРД. ДОЛЛ. США

По прогнозу Meticulous Research составит рост рынка к 2027 г.

26%

РОСТ РЫНКА В ГОД

Один из самых динамично развивающихся секторов цифрового здравоохранения

Самые перспективные тренды:

- ✓ Предписательная аналитика (рекомендательные системы)
- ✓ Облачные сервисы / SaaS
- ✓ Использование ЭМК как источника данных («топливо» аналитики)

Основные эффекты

- ✓ Сократить неэффективные потери благодаря прогнозированию событий, например пустые койки, потери просроченных лекарств и т.д.
- ✓ Оперативно перераспределять ресурсы при изменении нагрузки на МО, например при эпидемиях
- ✓ Выявлять пациентов высокого риска до момента манифеста ХНИЗ и тем самым существенно повысить эффективность профилактики
- ✓ Сократить ручную обработку и анализ больших данных, получать не текущие показатели (врачу, руководителю), а сразу опасные ситуации

Проблемы ЭМК как источника первичных данных

5

Состояние информатизации

- ✗ Свыше 40 компаний-разработчиков
- ✗ Отсутствие совместимости между различными ЭМК-системами
- ✗ Смешанная архитектура: централизованный и децентрализованный подходы
- ✗ Существенные различия уровня цифровой зрелости
- ✗ Отсутствие мотивации разработчиков к качественному выполнению проектов

Нормативное регулирование

- ✗ Отсутствие явного разрешения на сбор и обработку обезличенных медицинских данных
- ✗ Высокие требования к защите медицинской информации
- ✗ Большие сложности при подключении к ЕГИСЗ и региональным ГИС
- ✗ Большие сложности вывода программных медицинских изделий на рынок

Качество данных

- ✗ Низкое качество интерфейсов современных ЭМК-систем
- ✗ Шаблоны ведения медицинских протоколов
- ✗ Очень маленькое количество времени на ввод данных
- ✗ Отсутствие автоматического контроля качества, полноты и непротиворечивости данных со стороны ЭМК-системы

Причины сомнений в качестве данных ЭМК

6

Источники медицинской информации



Собранные со слов данные



Объективные данные



Данные с диагностического оборудования

Электронная медицинская карта Ручное внесение данных в интерфейс ЭМК



Автоматическая передача данных от оборудования в МИС

Особенности ведения ЭМК

- До 80% записей – текстовые документы, неструктурированные данные
- Чем более формализованная экранная форма для внесения информации – тем выше качество и ниже скорость заполнения



Формализованный первичный осмотр
Порядка 500-900 полей, справочники значений, ср. время заполнения – 30-40 мин.



Первичный осмотр на основе шаблона
Порядка 12 полей, шаблоны готовых текстов, ср. время заполнения 3-7 мин.

- Повторное использование (копирование) данных из предыдущих документов
- Отсутствие форматно-логического контроля полноты и достоверности данных
- Главный критерий ЭМК – удобство и скорость внесения информации



Обработка обезличенных ЭМК для оценки рисков развития заболеваний и выявления подозрений с помощью технологий искусственного интеллекта

WEBIOMED – это:



Первая и единственная Российская система поддержки принятия врачебных решений (СППВР), зарегистрированная как медицинское изделие (SaMD)



Анализ обезличенных данных пациентов, определение заболеваний и клинических состояний



Экспертное «второе мнение» для врача без ручного ввода данных и обработки информации во время приема



Внутренний контроль качества оказания медицинской помощи

98% Точность извлечения данных из ЭМК

50 прогнозных моделей

92% Точность прогнозов

25 NLP-моделей

Лидер среди российских ИИ-стартапов для здравоохранения *

- ✓ Победитель в номинации «Цифровые решения для здравоохранения» конкурса Аналитического центра при Правительстве РФ
- ✓ Цифровое решение, рекомендованное к внедрению и тиражированию в субъектах Российской Федерации
- ✓ Победитель в номинации «Прорыв года» премии Digital Health Awards 2020
- ✓ Победитель в номинации «Цифровая медицина» конкурса «Стартап-ралли 2020»
- ✓ Победитель конкурса Sanofi среди инновационных проектов в области цифрового здравоохранения
- ✓ Победитель конкурса «Технологии умной клиники» Сибирского государственного мед.университета

Источники «сырых» данных. Наши поставщики.

8

Региональные источники

- ✓ Ямало-ненецкий автономный округ
- ✓ Кировская область

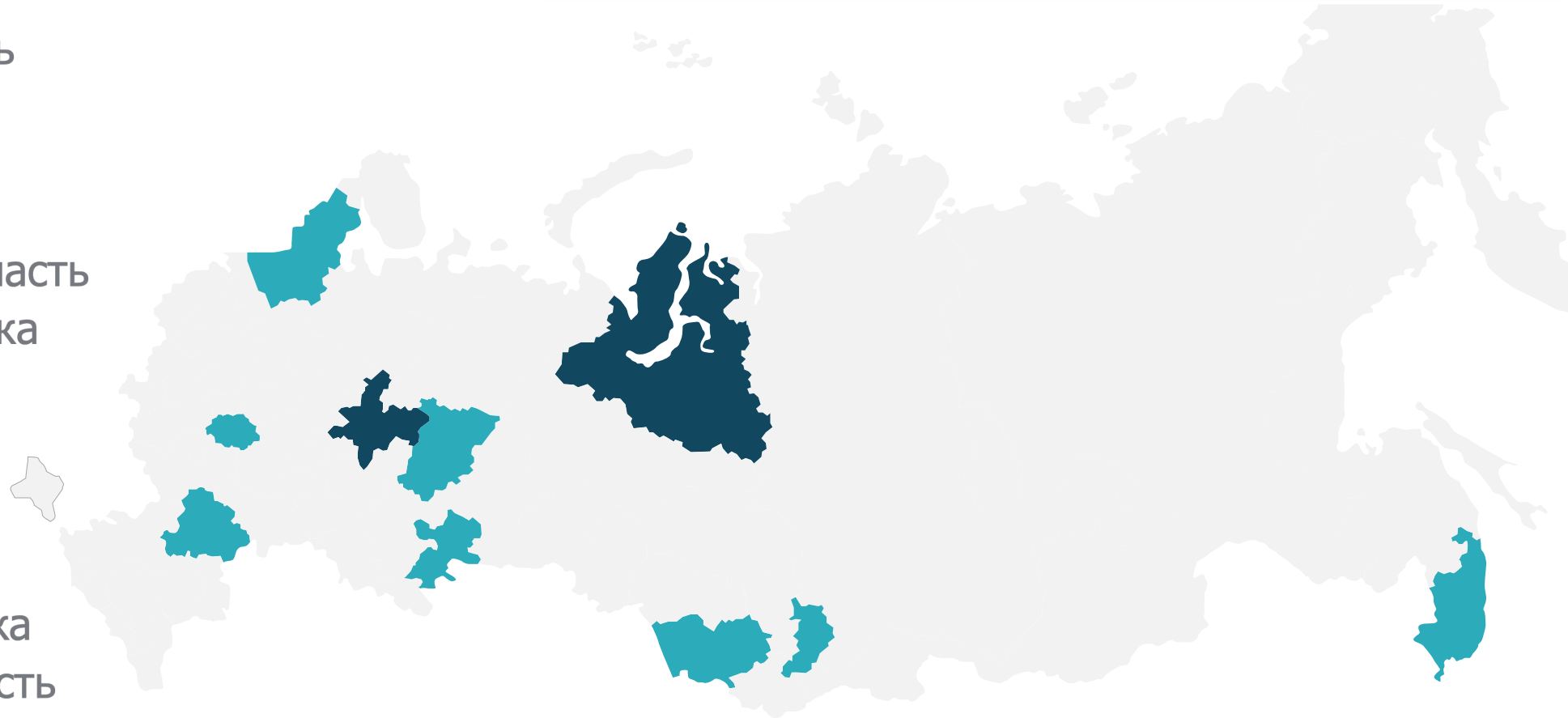
Отдельные МО

- ✓ Алтайский край
- ✓ Волгоградская область
- ✓ Карелия республика
- ✓ Пермский край
- ✓ Приморский край
- ✓ Рязанская область
- ✓ Санкт-Петербург
- ✓ Хакасия республика
- ✓ Челябинская область

87

Медицинских организаций

Загрузили в платформу обезличенные данные



Объем собранных «сырых» данных

9

2,8
млн

Пациентов

111
млн

Документов

19
млн

Амбулаторных
случаев

1,5
млн

Историй болезни

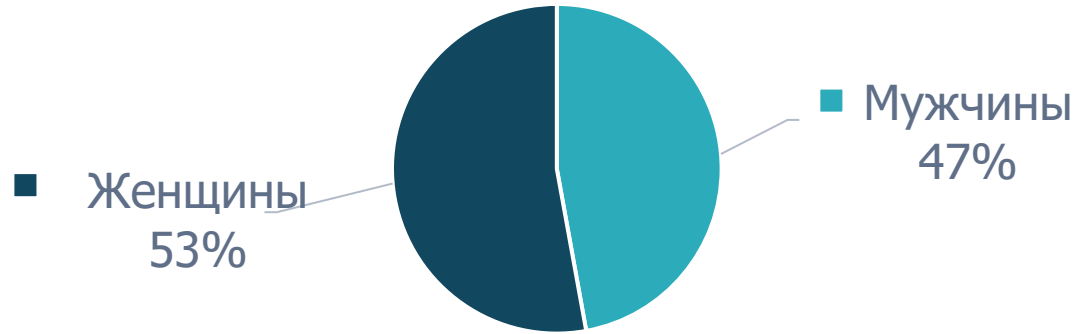
23

Вида СЭМДов

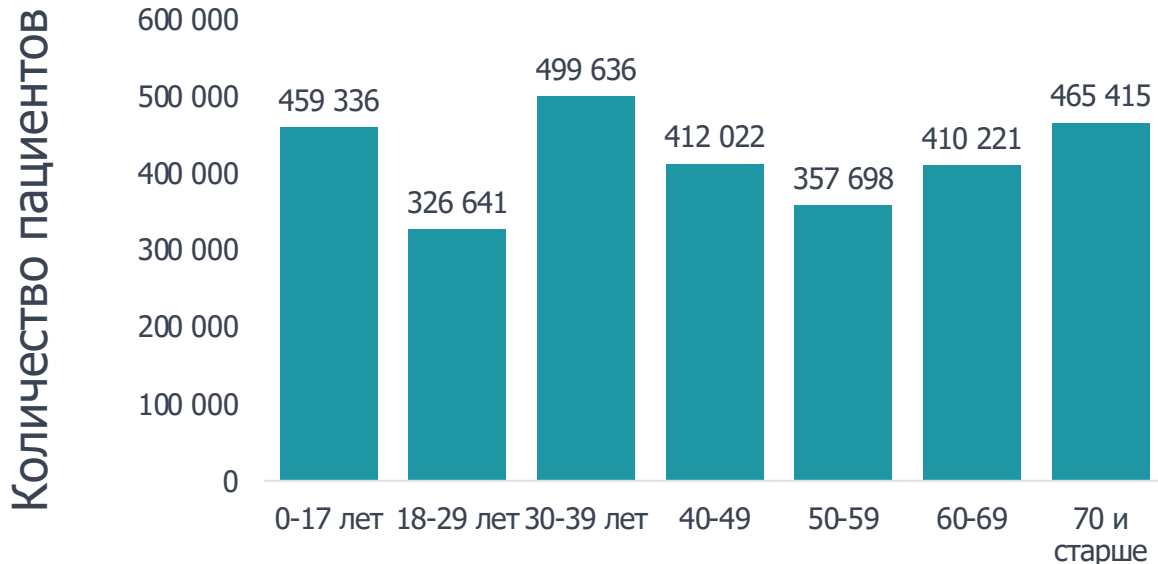
Тип документа	Доля, %
Протокол врачебного осмотра	49,51
Протокол лабораторного исследования	24,75
Электронный рецепт	10,59
Инструментальное исследование	9,05
Эпикриз по законченному случаю амбулаторный	2,25
Листок временной нетрудоспособности	1,89
Результат диспансеризации/профосмотра	0,62
Анкета для граждан в возрасте до 65 лет	0,49
Карта вызова скорой медицинской помощи	0,22
Анкета для профосмотра / диспансеризации	0,29
Медицинское свидетельство о смерти	0,16
Протокол операции	0,08
Анкета для граждан в возрасте 65 лет и старше	0,07
Эпикриз в стационаре выписной	0,02
Акт мед. освидетельствования на состояние опьянения	0,01
Анкета скрининга здоровья	0,01
Медицинское свидетельство о перинатальной смерти	0,00

Характеристика «сырых» данных в Webiomed

Распределение по полу



Распределение по возрасту



Статистика на 01.09.2021

возраст пациентов

20,7%

«Пустых пациентов»

В ЭМК не содержится ни одного медицинского документа

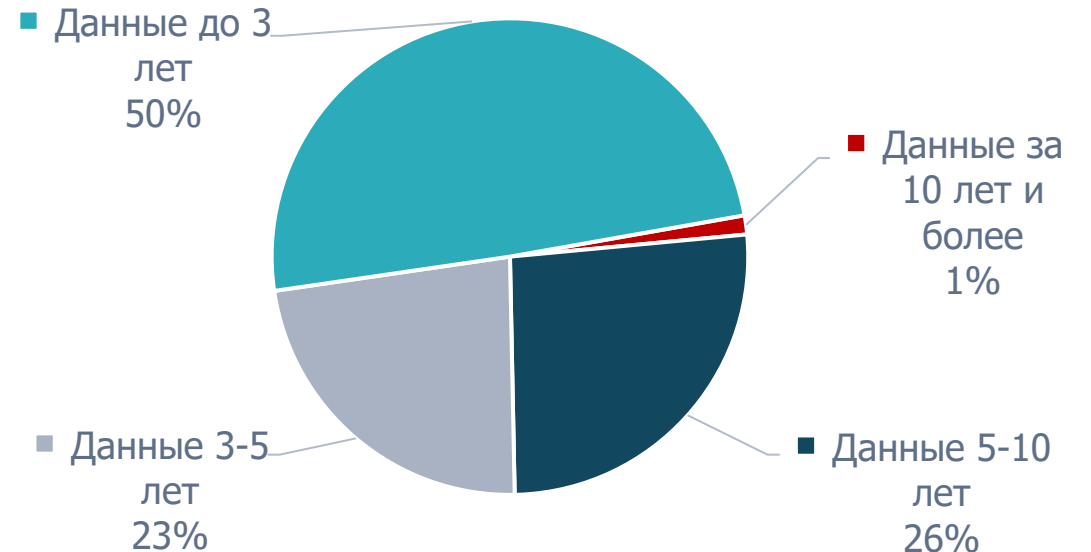
32,7%

Пациента с 1-10 документами
У таких пациент есть не более 10 медицинских документов в ЭМК

46,6%

Пациентов с > 10 документами
У этих пациентов есть минимально-необходимая для анализа информация

Глубина наличия данных



Характеристика извлеченных признаков

225
млн

Структурированных признаков

Было автоматически извлечено платформой из 111 млн. документов

560

Признаков

Поддерживает платформа в качестве видов данных, извлекаемых из ЭМК

7,5
млн

Ошибочных признаков

Было автоматически выявлено платформой на этапе ФЛК

3,4%

Доля ошибочных признаков

Которые были заблокированы для использования после извлечения

Название признака	Извлечено значений	Доля	В том числе ошибок	Доля
Систолическое АД	19 891 668	8,815	27 735	0,358
Диастолическое АД	19 711 467	8,735	36 438	0,471
Частота дыхания	19 612 957	8,692	95 171	1,229
ЧСС	17 147 068	7,599	217 338	2,806
Температура тела	12 975 919	5,750	32 149	0,415
Кашель	4 527 368	2,006	0	0,000
Утомляемость	4 345 898	1,926	0	0,000
СОЭ	3 898 311	1,728	81	0,001
Лихорадка	3 774 808	1,673	0	0,000
Лейкоциты	3 772 882	1,672	357	0,005

Не вся ЭМК – это «плохие» данные



Источник: результаты анализ данных платформы Webiomed

Выводы о пригодности ЭМК для аналитики

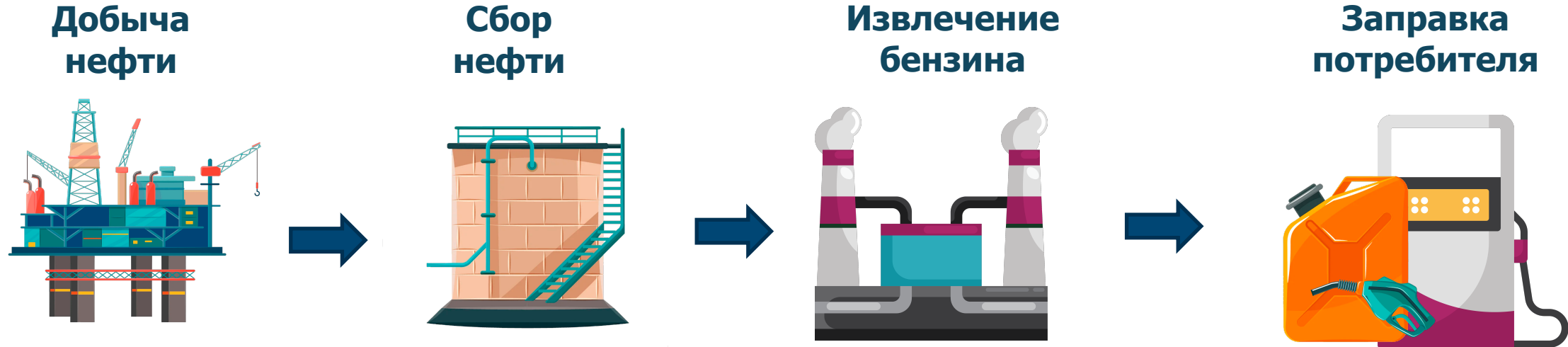
- ✓ В среднем **80%** ЭМК пациентов содержит минимально-необходимую медицинскую информацию
- ✓ До **50%** ЭМК содержат данные не более чем 3 последних года
- ✓ Лишь **26%** ЭМК содержат данные наблюдения от 5 до 10 лет
- ✓ Около **50%** ЭМК – это протоколы врачебных осмотров
- ✓ Около **33%** ЭМК– протокол лабораторного и инструментального обследования
- ✓ Около **3,4%** извлеченных из ЭМК признаков являются ошибочными
- ✓ Порядка **95%** ЭМК можно доверять при контроле качества

Данным ЭМК **можно верить и использовать.**
Но форматно-логический контроль (ФЛК) **критически важен!**

Схема использования ЭМК для СППУР

14

Формирование качественных наборов данных (data set) из электронных медицинских карт (ЭМК) можно сравнить с производством высоко-октанового бензина из сырой нефти



Системы
ведения ЭМК

Система
централизованного
сбора «сырых»
данных
(Data Lake)

Система
извлечения
данных
(NLP-сервис)

Цифровой
двойник
пациента, МО

... и перейти к системам поддержки принятия управленческих решений (СППУР) на основе ЭМК

- ✓ ЭМК – источник первичных данных
- ✓ Автоматический сбор и подготовка
- ✓ Искусственный интеллект как технология анализа
- ✓ ВІ как визуализация





<https://webiomed.ai/>



ВКонтакте

<https://vk.com/webiomed>



Facebook

<https://www.facebook.com/webiomed/>



Twitter

<https://twitter.com/webiomed>



Telegram

<https://t.me/webiomed>



YouTube

<https://www.youtube.com/>

Спасибо за внимание!