



# Возможности ВІ для анализа ИЭМК

---

08.06.2017

# РС ЕГИСЗ: множество источников данных

Хранилища медицинских данных:

- Региональная ИЭМК (МИСы, ЛИСы, РИСы и др.)
- Региональные нозологические регистры (Онко, ОКС и др.)
- Специализированные региональные сервисы (лекарственное обеспечение, ИПРА и др.)

Учетные системы:

- Кадровые
- Бухгалтерские

В каждой системе свои справочники и для анализа данных их необходимо связывать.

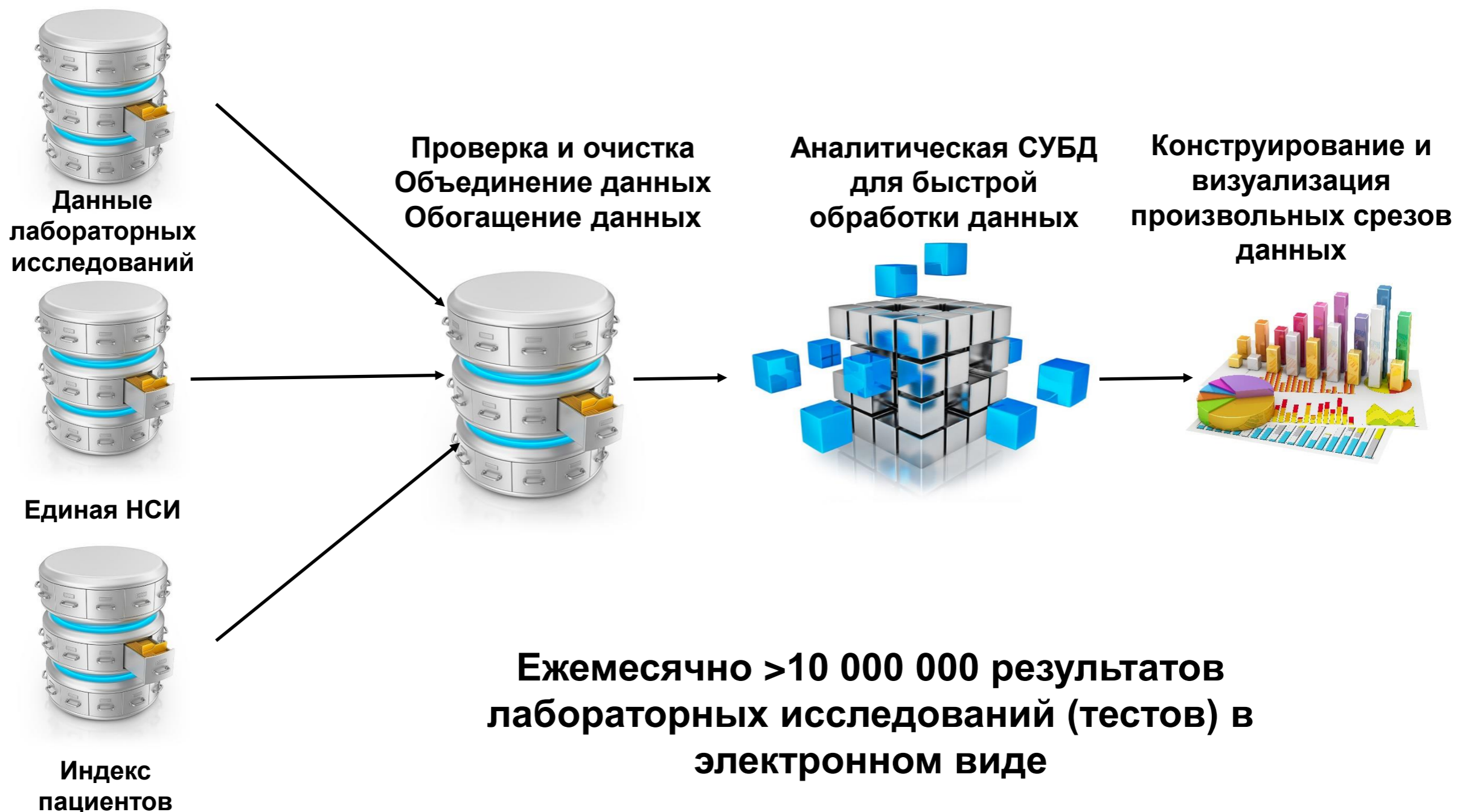
# Типовые сложности использования отчетных систем

- «Данных собираем много, а на запрос Минздрава не ответить»
- «Дайте конструктор – мы сами сформируем нужный отчет»
- «А как рассчиталась эта цифра?»
- «Почему так долго формируется отчет?»
- Разные модули РС ЕГИСЗ -> разные отчетные системы -> не связанные отчеты

Мало реализовать веб-сервисы и сложить данные в хранилище.

Для получения знаний нужно иметь возможность изучать информацию в различных срезах.

# Данные – информация - знания



# ETL: Извлечение, преобразование, загрузка

## Проверка данных и очистка:

- Пустые поля: пробел, «», Null, табуляция и т.п.
- Ошибки кодировки для строковых параметров

## Объединение данных из ОДЛИ, НСИ, Индекса пациентов:

- Коды медицинских организаций и КДЛ
- Идентификаторы пациентов
- Коды услуг и коды тестов ЛАТЕУС
- Версионность справочников (!)

## Обогащение данных:

- Предрасчитанные значения: возраст, отклонение от нормы

**Netrika.**

# Быстрая работа с большими данными

## Например, ClickHouse:

- Столбцовая СУБД для работы с большими данными
- Свободно распространяемое ПО
- Разработано Яндексом для второй по величине системы web-аналитики в мире

## Выборка результатов лабораторных исследований:

- детализация по пациентам
- 5 онкомаркеров, с отклонением от нормы

СУБД	Период	Время
PostgreSQL	Месяц	1-2 часа
ClickHouse	Месяц	1-2 мин

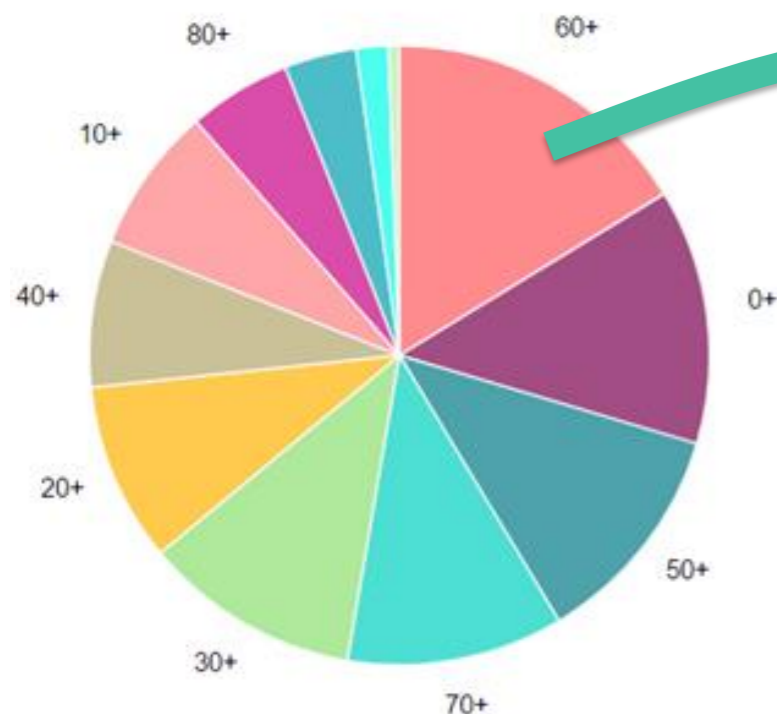
# Новые возможности для анализа информации

Например, **Superset**:

- Быстрая визуализация любых срезов детальных данных
- Детализация любой цифры: до врача, пациента, оказанной услуги... и еще детальнее!
- Конструктор отчетов: всё в руках пользователя
- Связывание аналитических форм между собой

# Пример связанных аналитических форм

Количество результатов (с отклонениями), по возрастам



Дата рождения	Возраст	Врач	Код теста	Название теста	Результат	Норма
1947-08-26	69	Ольга Евгеньевна Л [REDACTED]	69405-9	Скорость клубочковой фильтрации	53.8	'''
1950-10-21	66	Татьяна Григорьевна В [REDACTED]	17856-6	Гемоглобин гликированный (A1c)	5.68	6.5-12
1955-09-08	61	Дарья Александровна У [REDACTED]	17856-6	Гемоглобин гликированный (A1c)	5.9	6.5-12



# Перспективные направления анализа данных ИЭМК

## В части лабораторных исследований:

- Мониторинг качества лабораторной диагностики на основе использования статистических методов (среднее значение нормы)
- Контроль необоснованных повторных назначений лабораторных исследований
- Исследование частоты выявления патологии по районам Санкт –Петербурга (району проживания пациента)
- Контроль маршрутизации лабораторных исследований
- Контроль сроков диагностики пациентов с подозрением на онкологическое заболевание простаты

# Спасибо за внимание!

Игорь Башков

Руководитель отдела продаж и развития  
направления «Здравоохранение» компании  
«Нетрика»

[sales@netrika.ru](mailto:sales@netrika.ru)  
+7-906-269-03-34