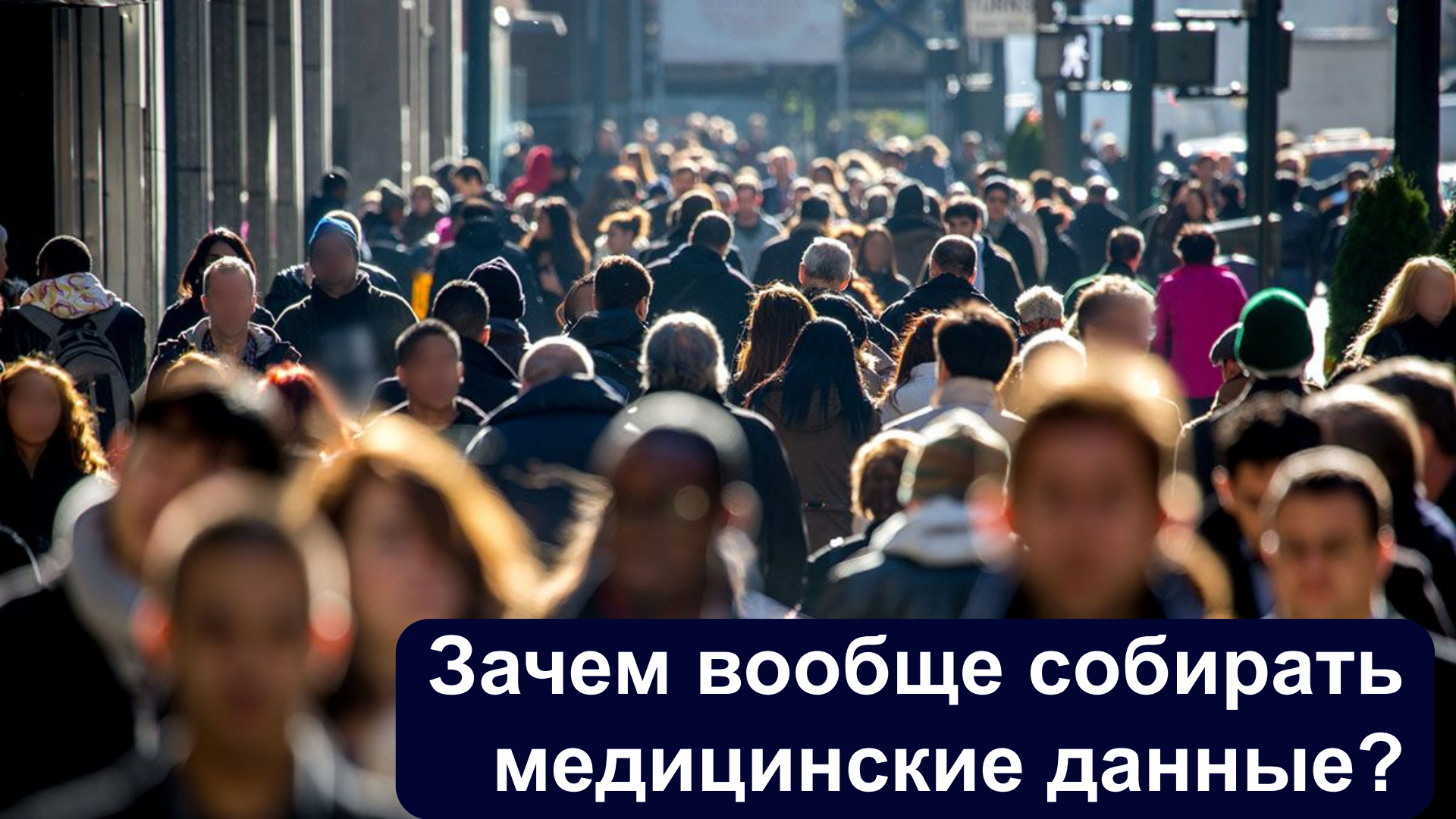
Проблемы организации сбора данных в офтальмологии

Андрей Бурсов
Врач-офтальмолог
Советник по цифровой медицине
ИСП РАН

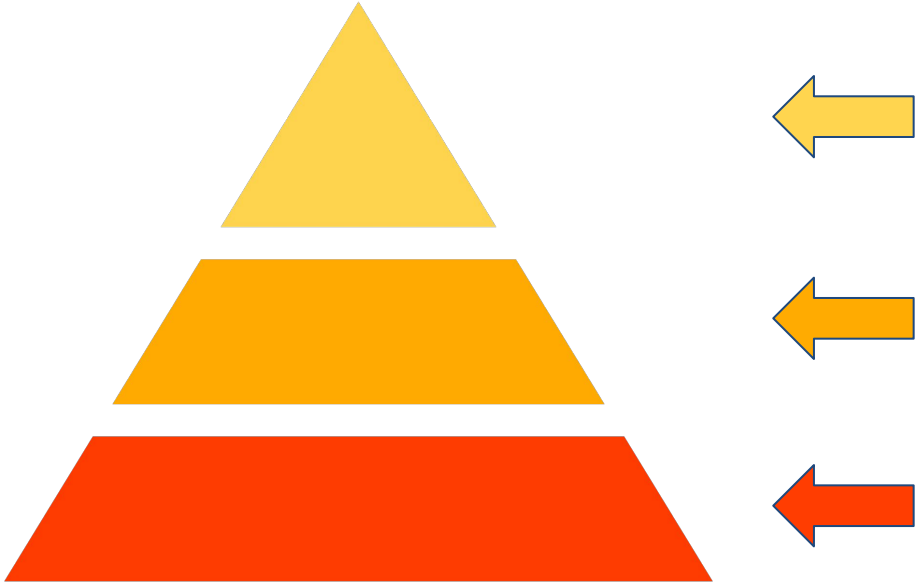
“Данные - новая нефть”

Джек Ма, создатель компании Alibaba



Зачем вообще собирать медицинские данные?

“Пирамида” искусственного интеллекта

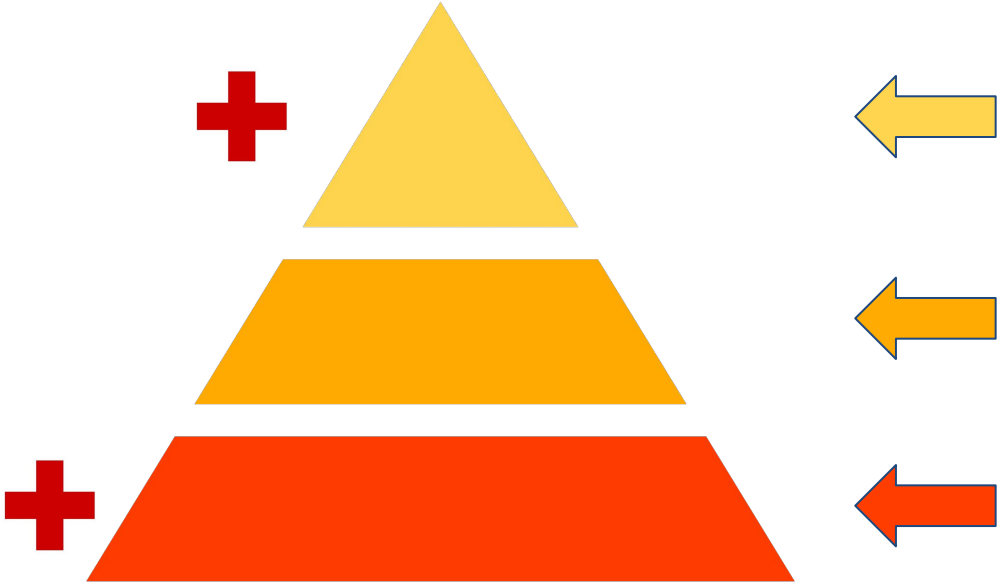


СППВР

Машинное обучение

Данные

“Пирамида” искусственного интеллекта



СППВР

Машинное обучение

Данные

Что нужно IT?

Компетенции

Ваши знания

Ресурсы

Ваше время
Ваши данные

Источники данных

Существующие
Создаваемые

МИС

Приборы

Открытые источники

Личная коллекция

Карта рынка цифровой медицины

Авторы карты

EverCare

СКАЧАТЬ ПЕЧАТНУЮ ВЕРСИЮ

Версия 1.3 от 15 мая 2021

Компании по категориям

72 МИС

Медицинские информационные системы

3

Большие данные в медицине

17

Искусственный интеллект в медицине

12

Персональная генетика

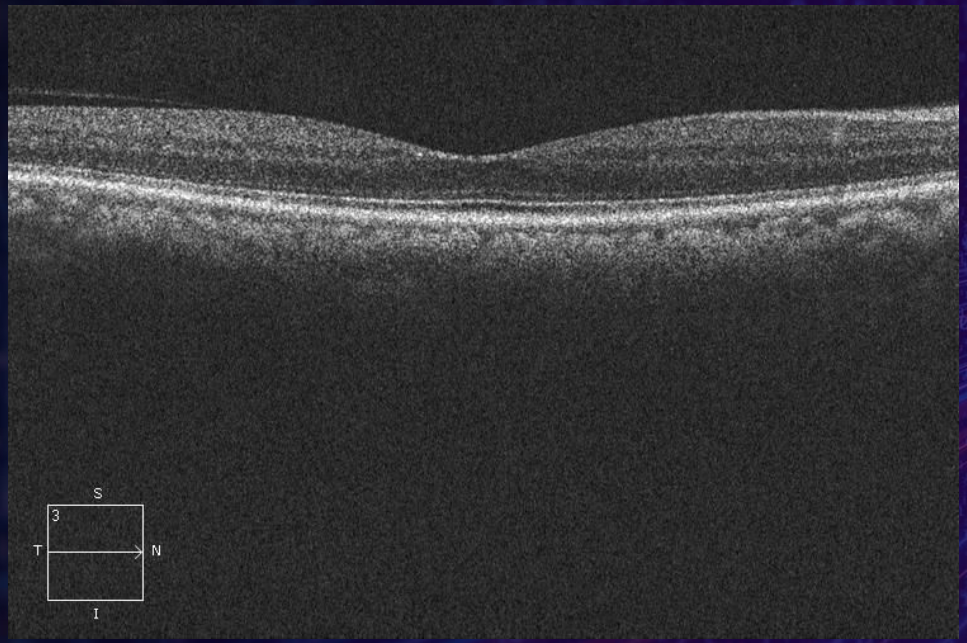
Сколько данных достаточно?*

100 / 1000 / 1 000 000+

*на одну стадию патологии

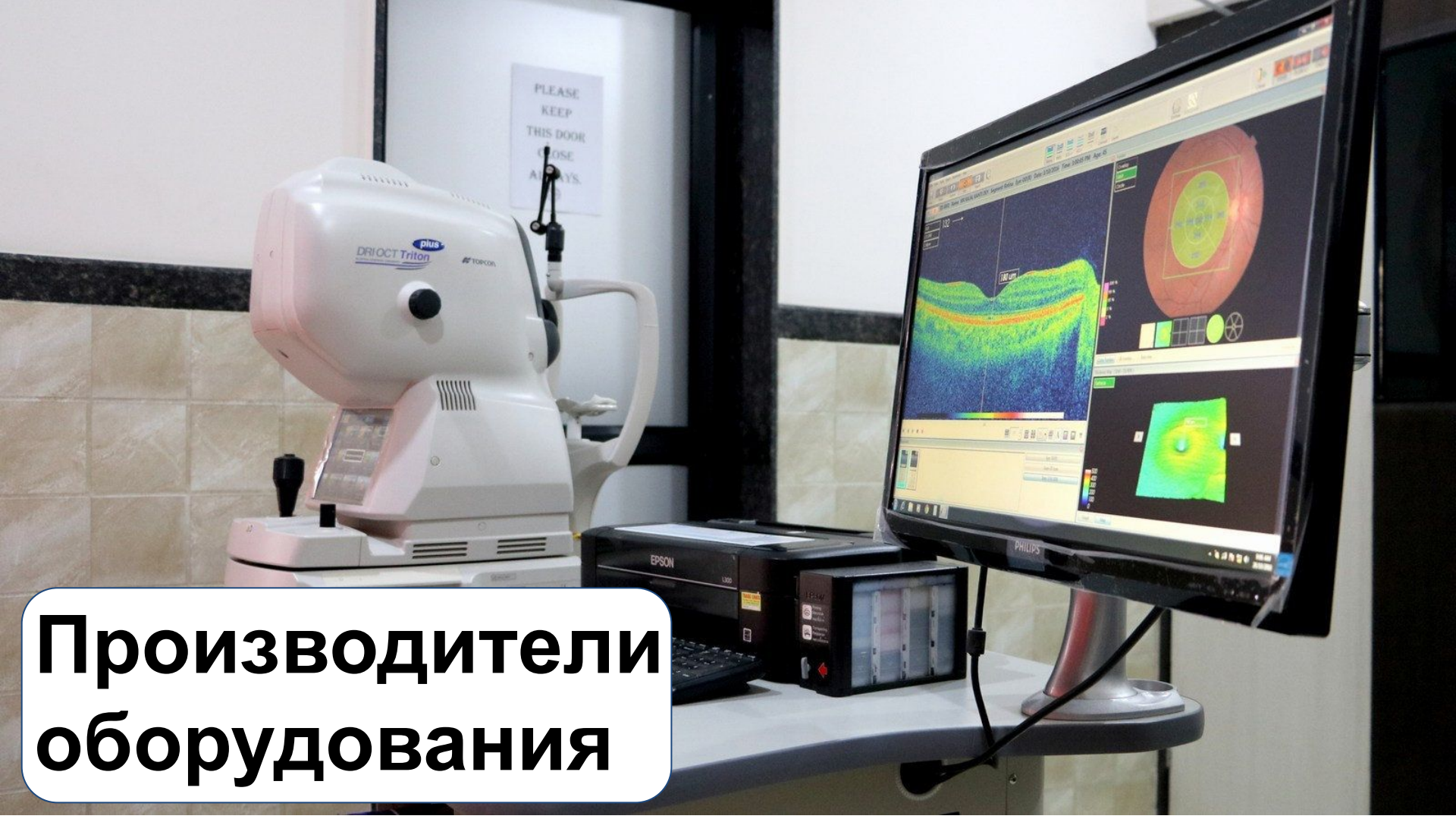
Какие офтальмологические данные сейчас в тренде?

Какие офтальмологические данные сейчас в тренде?



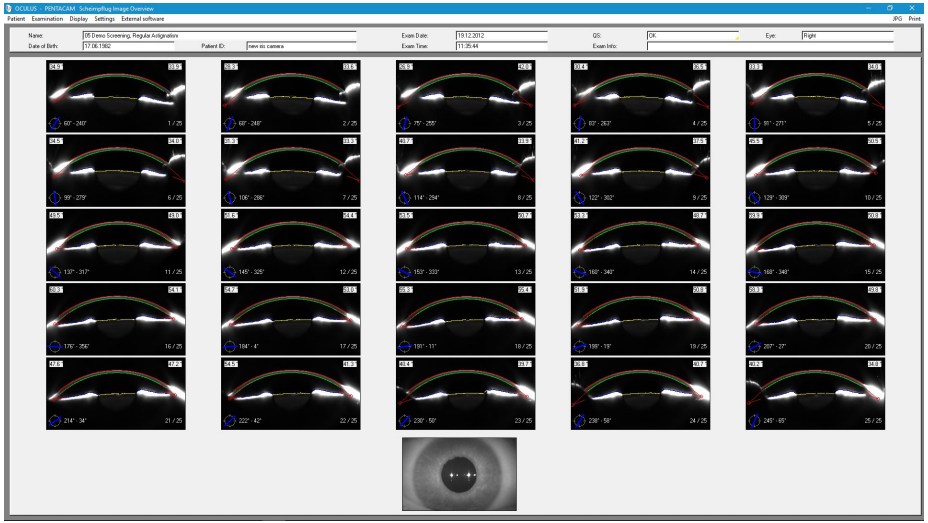
ОСТ

Фоторегистрация (Fundus)



Производители оборудования

Pentacam



1. Гридин В.Н., Труфанов М.И., Бурсов А.И., Виноградов В.А. Устройство для определения геометрических характеристик роговицы глаза // Патент на полезную модель RU 163169 U1, 10.07.2016. Заявка № 2016103084/08 от 01.02.2016.
2. Гридин В.Н., Труфанов М.И., Бурсов А.И., Виноградов В.А. Оптико-электронное устройство для диагностики офтальмологических заболеваний // Патент на полезную модель RU 184949 U1, 15.11.2018. Заявка № 2018103965 от 02.02.2018.

У офтальмологов есть огромная персональная нестандартизованная распределенная база данных глазной патологии

Что хранят открытые данные?

67 Results Sort by: Relevancy

Dataset

Diabetic Retinopathy 224×224 Gaussian Filtered

by Sovit Ranjan Rath

2 years ago • 448 MB • 117

Diabetic Retinopathy 224×224 Gaussian Filtered

Dataset

Diabetic Retinopathy Arranged

by Aman Neo

9 months ago • 1 GB • 21

Diabetic Retinopathy Arranged

Dataset

Diabetic Retinopathy 2015 Data Colored Resized

by Sovit Ranjan Rath

2 years ago • 2 GB • 22

Diabetic Retinopathy 2015 Data Colored Resized

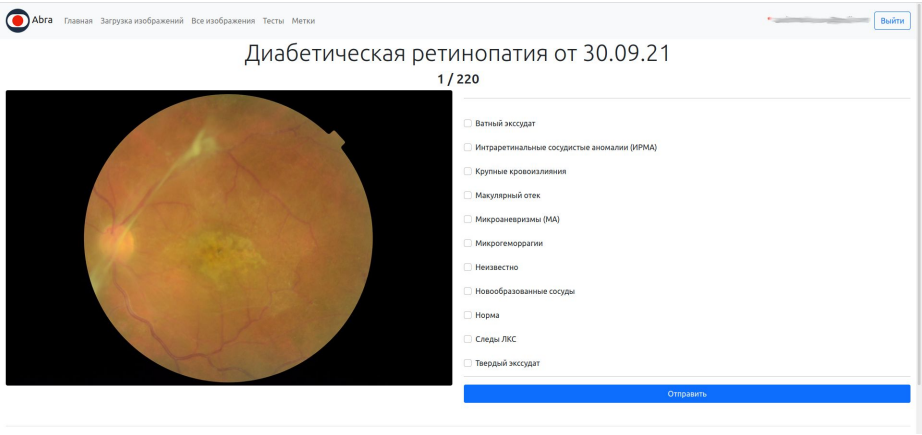
1 Total rows loaded: 298508

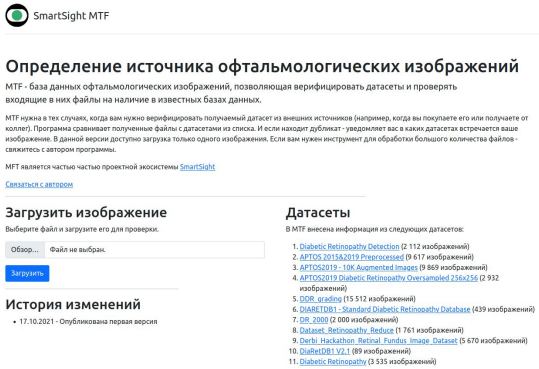
hash	COUNT(*)
1dd6600d01dd0ff1e42ce840ffa501ae32b346745ff4c4f9296e62a2e9632f88	14
007833ca3020cc4945b8300ed8d74d7a255aa0effe9f8f0bbf0a71479d946546	11
075af32204aa6f3ac094d181ee472391c4542d0251034e0c054928e447fcd89f	11
08a4d41cac57a08c3d97dfd10d40554a48b93d4d912bfc4a6cb3d14f97d5ae4a	11
162cad82834dc12c19d44518ce019bc7020812079ece8b8e81699ec5c94f366b	11
186a9dadbc7452eb3f80de76e773d9b0dc5a2dab9f3c59a84990a09b7f6b3abf	11
235da5685451a3f5dcd21a1c2543388730bbfac7f2ebe4f4fda509d1423ce0dd	11
25a33cc293d7f0ca4579f26b6dbbfaa7cad27dc20bfbee0fc7114968f44115fd	11
388c258d7288d05cc35ce23476d6b9e461fbcf782e18b7cae8d7e635bf9b54a5	11
399c283617595d2b19126f5cf9df8d0d8bd755dcd20bc60aa6145ab408913714	11
585e8833f811facdea38b4aafcf357f99c6b60ea4b546de2d95f1f71d80474ed	11
5a3f78bd971bfa888529106e0f05788ac70491faaf03c8b864a3333e43961381	11
63e253f2934ad6abb3030f2e01b968f87c34c4138824a872cab00a802aa46aad	11
810606d49e4cd3fa93bd25209fcea371203b6bd59707a9ebcb45db97cbed9b89	11
8ddf5ea2f79222ec3c1f445e9075fd0f06a039aa92d28eba9ed6a16c23715cdb	11
aa203da2bbbd55e2ddcabbfdc0fa59ad2248488122793c58d5a97d6027b869e0	11
b6c623b4d25321e592b6f35949be2040f2efcb1f15019efedfa96454ff970d23	11

>500 Gb kaggle.com
1 371 922 изображений

Повторяемость данных - 21,7%

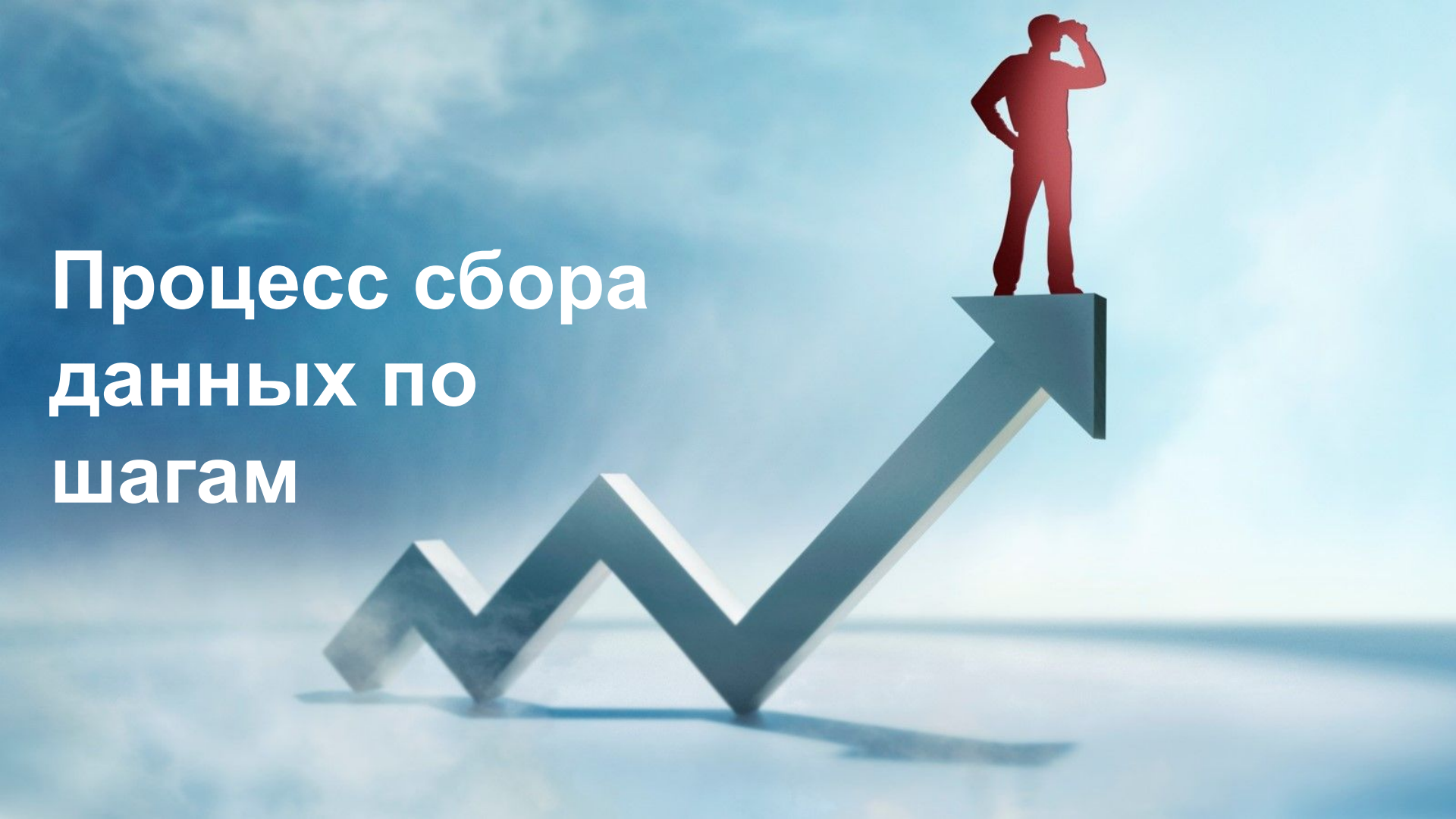
SmartSight





1. Разметка цветных фотографий глазного дна улучшает распознавание макулярной патологии с помощью глубокого обучения, Х. П. Тахчиди, П. В. Глизница, С. Н. Светозарский, А. И. Бурсов, К. А. Шустерзон // Вестник РГМУ. 2021 / 04. https://doi.org/10.24075/vrgmu.2021.040
2. Бурсов А.И., Глизница П.В., Светозарский С.Н. “Прототип интеллектуальной системы диагностики глазного дна и структур переднего отрезка глаза для мобильной фундус-камеры” // Свидетельство о регистрации программы для ЭВМ RU 2021666454, 14.10.2021. Заявка № 2021663283 от 25.08.2021
3. Тахчиди Х.П., Качалина Г.Ф., Тахчиди Н.Х., Бурсов А.И., Маноян Р.А., Абрамова О.И., Глизница П.В. Адаптер для крепления смартфона на щелевой лампе // Патент на полезную модель RU182293 U1, 13.04.2018, заявка №2018113408 от 13.04.2018

Процесс сбора данных по шагам



Постановка задачи?

КАК И ЗАЧЕМ?

Где и как хранить?

Кто эксперты?

Постановка задачи?

Бюджет, сроки?

Откуда данные?

Как размечать?

Как размечают данные?

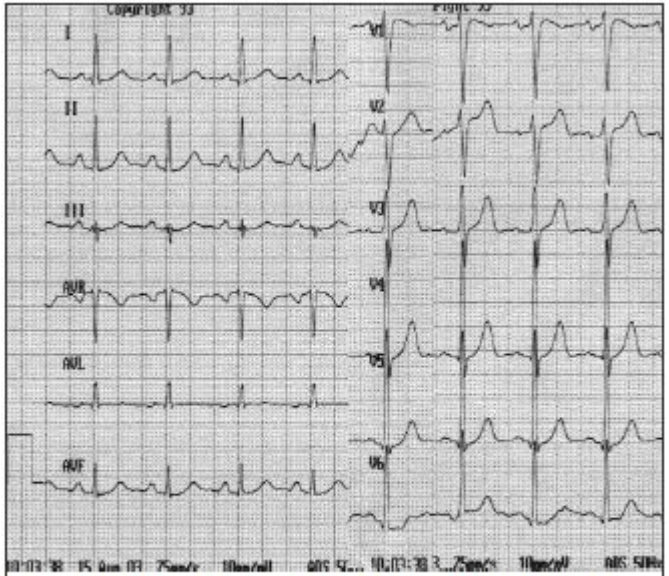
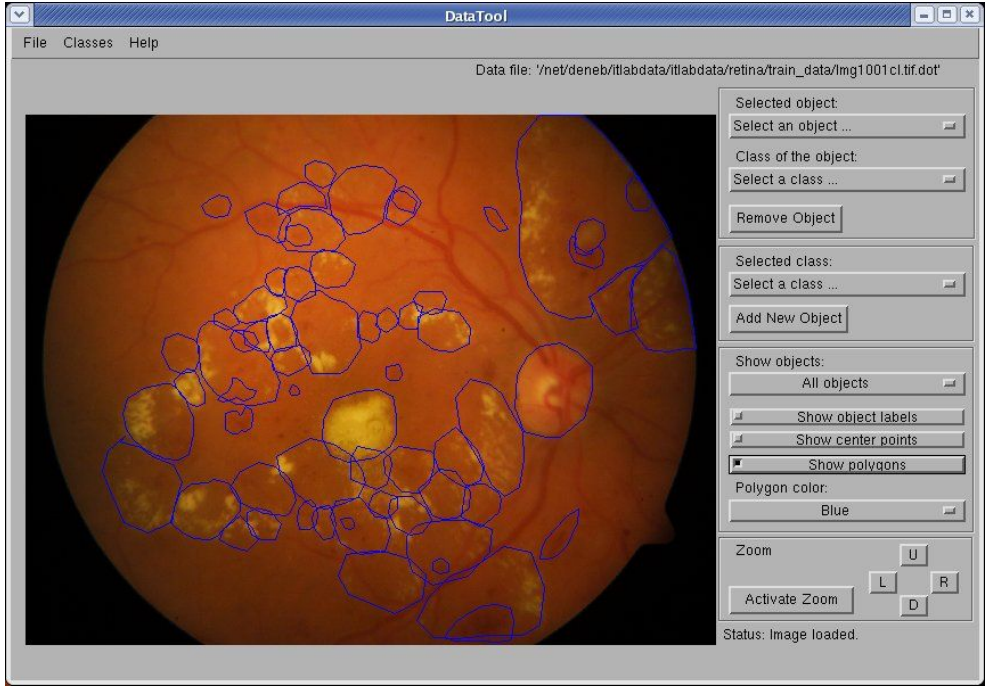


Рис. 1. ЭКГ при синусовой тахикардии. (+)P II, III, aVF. Зубец P, интервал PQ, R-R – нормальные. Правильный ритм с ЧСС 102 в 1 мин



Тем временем в проекте…

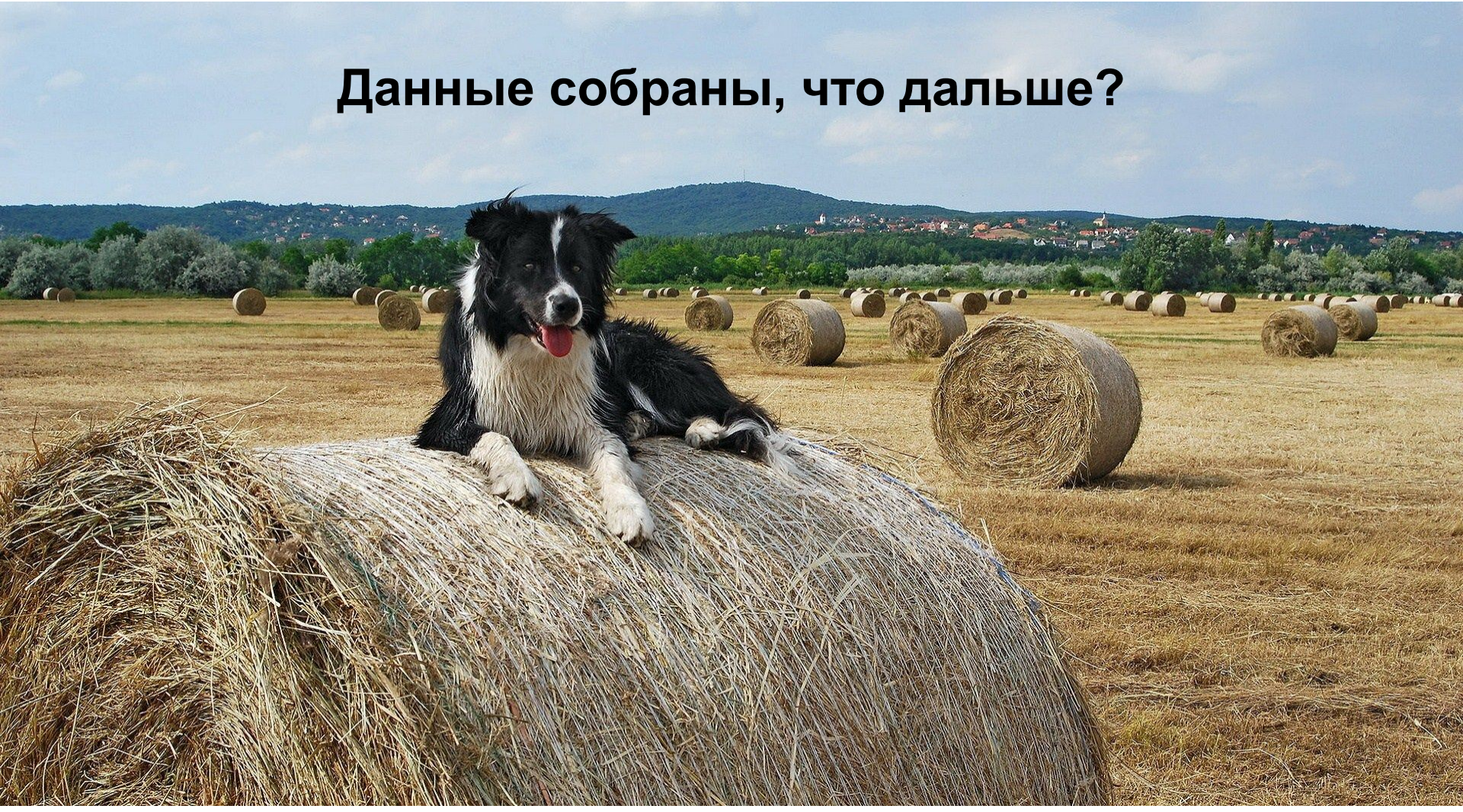
- **<u>Планирование разметки</u>**
- **<u>Нагрузка на врачей и оплата</u>**
- **Технические сложности**
- **Распределенная команда**
- **Верификация**
- **Человеческий фактор**

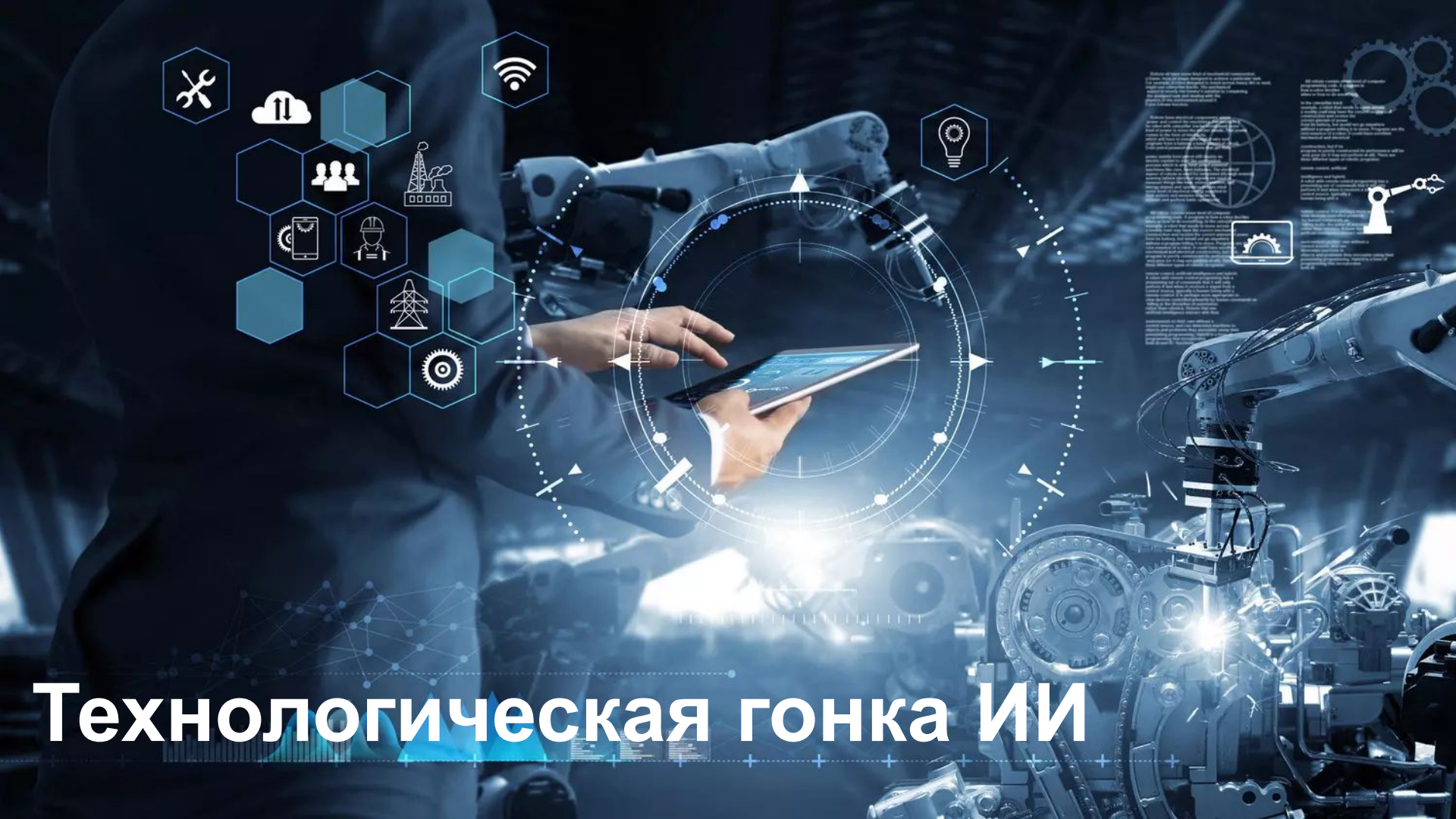
Масштабирование

Масштабирование

Совет: база знаний очень помогает

Данные собраны, что дальше?





Технологическая гонка ИИ

Спасибо за внимание!

Проблемы организации сбора данных
в офтальмологии

https://bit.ly/3E1Df63

Андрей Бурсов

Советник по цифровой медицине ИСП РАН
bursov@ispras.ru