

УНИВЕРСИТЕТ ИТМО

Поддержка принятия решений
в городском здравоохранении:
от больших данных – к предсказательному
моделированию

Александр Валерьевич Бухановский

Стратегические проекты развития Университета ИТМО

1. Фокусировка в приоритетных направлениях развития Университета
2. Дальний горизонт сохранения конкурентоспособности
3. Коллаборативный характер (формирование устойчивой экосистемы вокруг проекта)
4. Междисциплинарность, «игра» на симбиозе направлений
5. Комплексность, покрытие жизненного цикла продукции
6. Международная узнаваемость, позиции на зарубежных рынках
7. Локализация: интересы города/региона
8. Трансформация образовательного процесса

Длительность проектов: 4-5 лет

Себестоимость проектов: 500-800 млн. руб.

Стратегический проект «Urban Health Decision Support: from Big Data – to Predictive Modeling»

Эффективное управление здравоохранением на уровне большого города с использованием предсказательных моделей, построенных на больших данных

Основа управления: определение управляющего контура и задание целевой функции (оценка эффекта)

Computer Science

Skoltech

Skolkovo Institute of Science and Technology

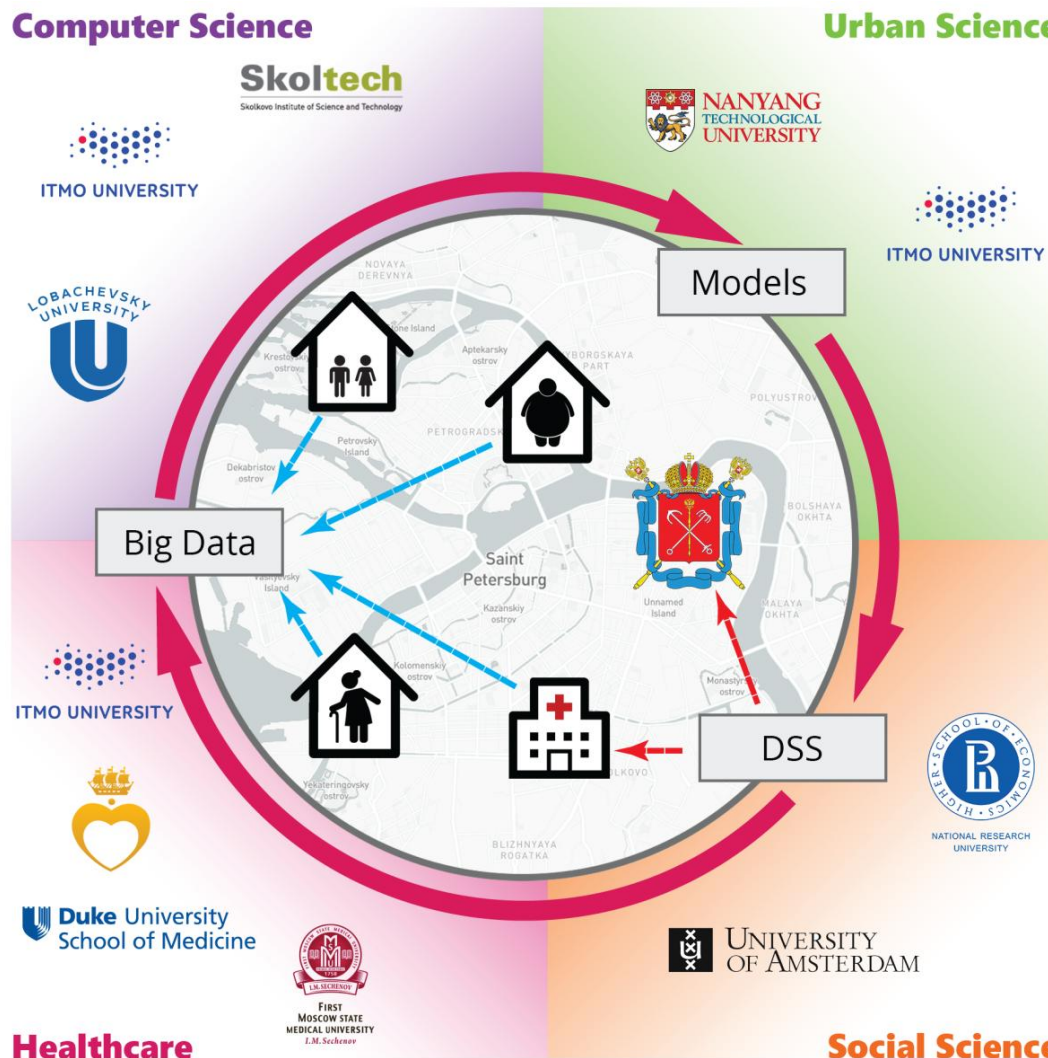


Healthcare

Urban Science



Social Science



Цели стратегического проекта

1. Обоснование и разработка теоретических основ эффективной организации и управления процессами городского здравоохранения на основе ценностно-ориентированного подхода (value-based approach, VBA)
2. Создание информационной технологии для оценки, прогнозирования и оптимизации качества жизни и здоровья горожан на основе количественных показателей в рамках VBA с учетом различных масштабов изменчивости городской среды
3. Развитие информационно-вычислительной инфраструктуры для поддержки принятия решений в области городского здравоохранения, обеспечивающих повышение качества жизни и здоровья населения мегаполисов

Современное здравоохранение: Volume vs. Value



Не в то
время, не в
то место!



Специфические городские группы – потребители VBA:

- 1) Пожилые люди (в том числе, работающие пенсионеры)
- 2) Молодые люди в группах риска (малоподвижные менеджеры)
- 3) Трудоспособные люди, страдающие психиатрическими и психосоматическими заболеваниями
- 4) Дети и подростки, растущие в городской среде

Ценность: проблема измеримости

$$\text{Ценность} = \text{Результат} / \text{Цена}$$

$$\text{Результат} = \left\{ \begin{array}{l} \text{Клинический эффект - для врача} \\ \text{Качество жизни (QoL) – для пациента} \\ \text{Экономический эффект – для страны} \end{array} \right.$$

$$\text{QoL} = \left\{ \begin{array}{l} \text{Физиологическое благополучие} \\ \text{Социальное благополучие} \\ \text{Психическое (душевное) благополучие} \end{array} \right.$$

QoL – субъективная популяционная характеристика



Ключевая идея проекта - смешанные данные:

- что не измерить, то можно оценить; что не оценить – то можно смоделировать...

**Физиологическое
благополучие**

=

- Hard Data (Медицинские ИС)
- Hard Data (Носимые устройства)
- Soft Data (Социальные медиа)

**Социальное
благополучие**

=

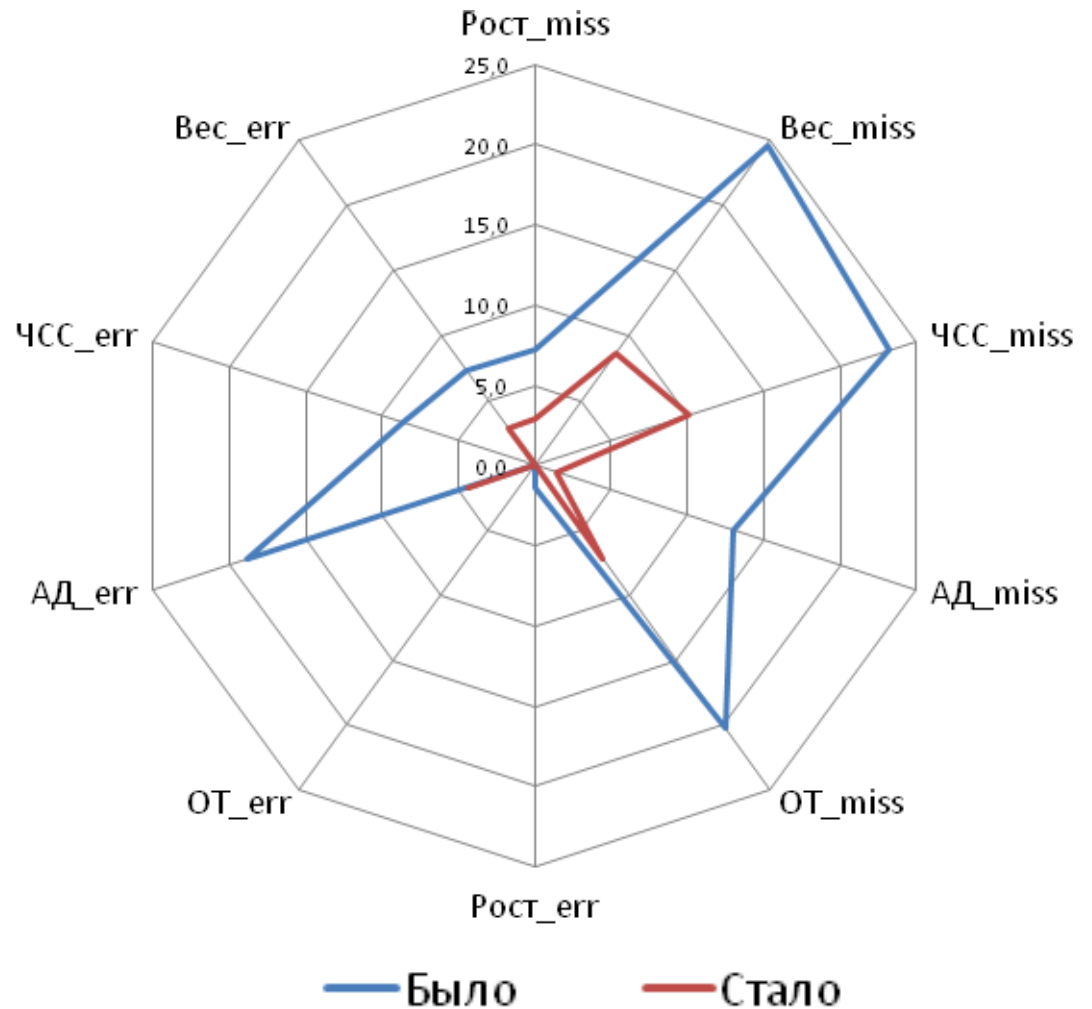
- Hard Data (Носимые устройства, с геопозиционированием)
- Synthetic Data (Модели городской среды)

**Душевное
благополучие =**

- Synthetic Data (модели влияния городской среды)
- Soft Data (Социальные медиа)

Проблемы использования Hard Big Data МИС

- 1) Слабоформализованное представление данных (текстовые поля, заполняемые с «естественной» грамотностью
- 2) Пропуски и засорения при вводе формализованных данных
- 3) Рассинхронизация временных характеристик данных при коллективной работе
- 4) Умышленные искажения

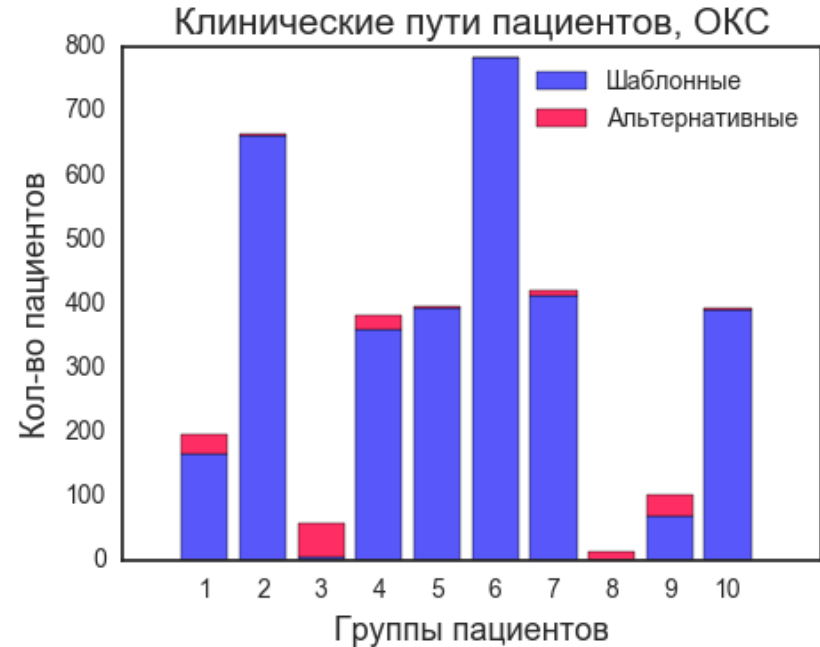


Снижение ошибок ввода первичных данных в МИС за счет использования интеллектуального интерфейса

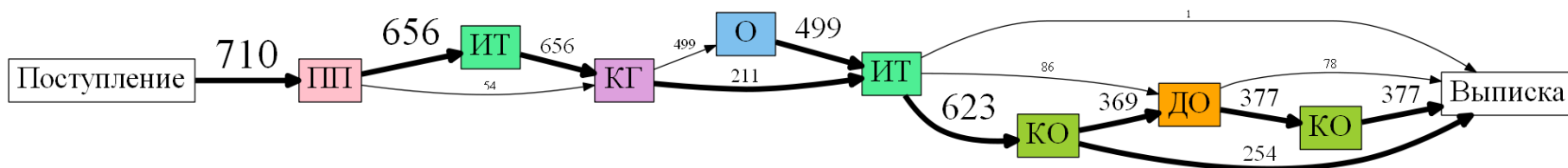


Восстановление структуры данных – модель клинических эпизодов

Классификатор	Точность
QdiscriminantAnalysis	74 %
DecisionTreeClassifier	75 %
AdaBoostClassifier	78 %
KNeighborsClassifier	81 %
RandomForestClassifier	81 %



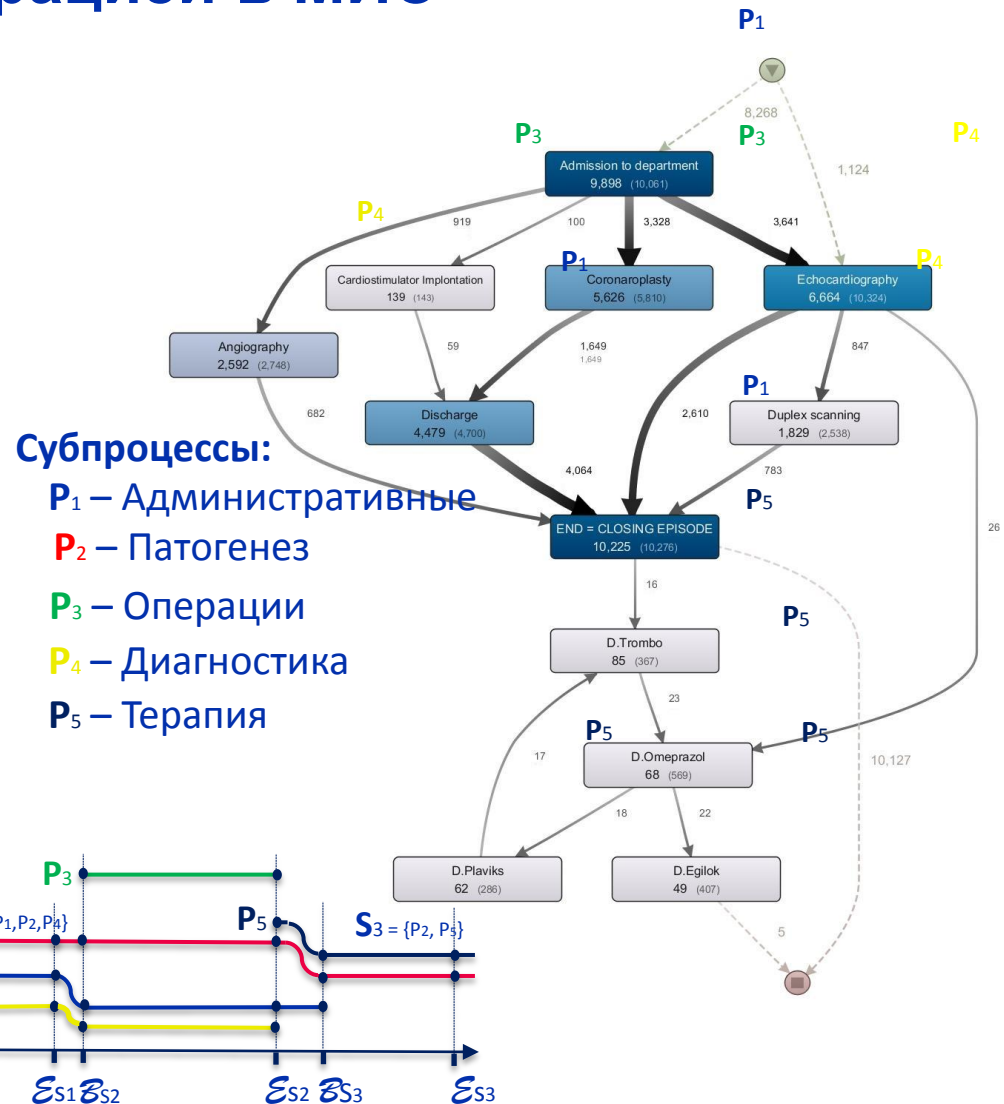
Автоматическое исправление ошибок в записях (до 90% - восстановление порядка записей; до 10% - определение ошибочных записей)



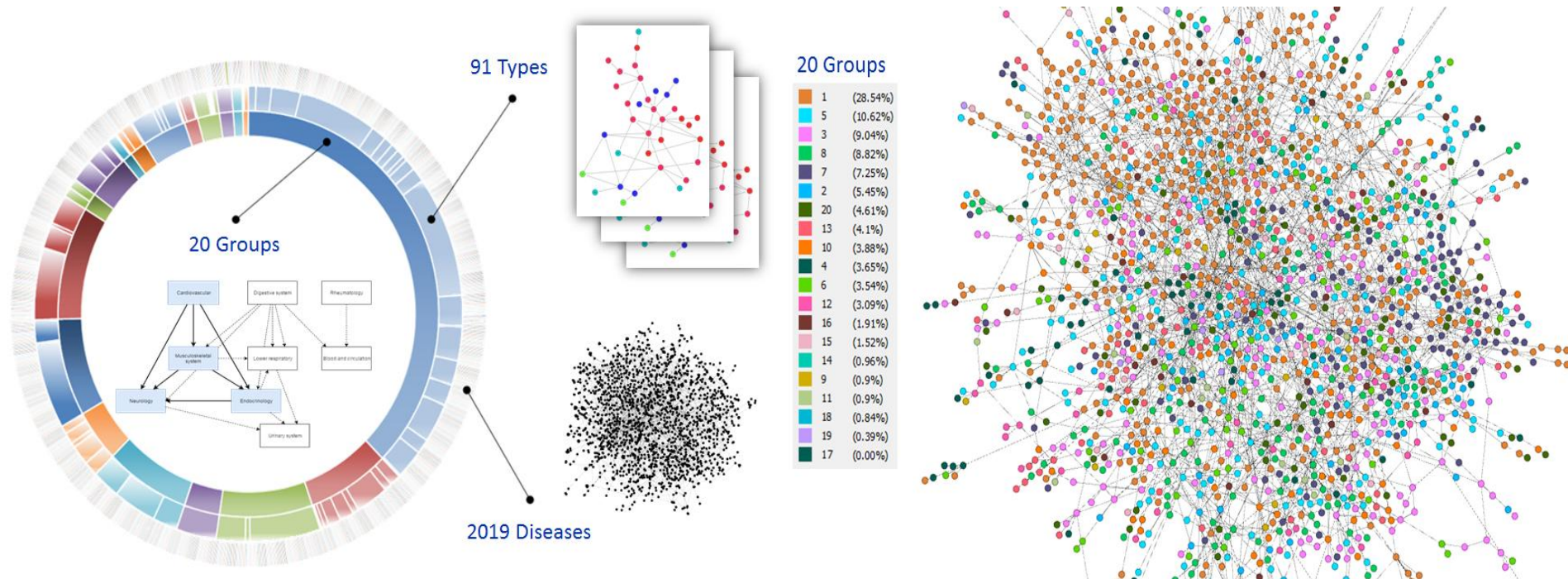


Идентификация комплексных процессов по ЭМК с интеграцией в МИС

- 1) Формализация структуры процессов методами Process Mining
- 2) Идентификация параметров процессов в рамках заданной структуры
- 3) Описание неоднородности в пространстве параметров на основе методов классификации
- 4) Исследование причин неоднородности и их интерпретация



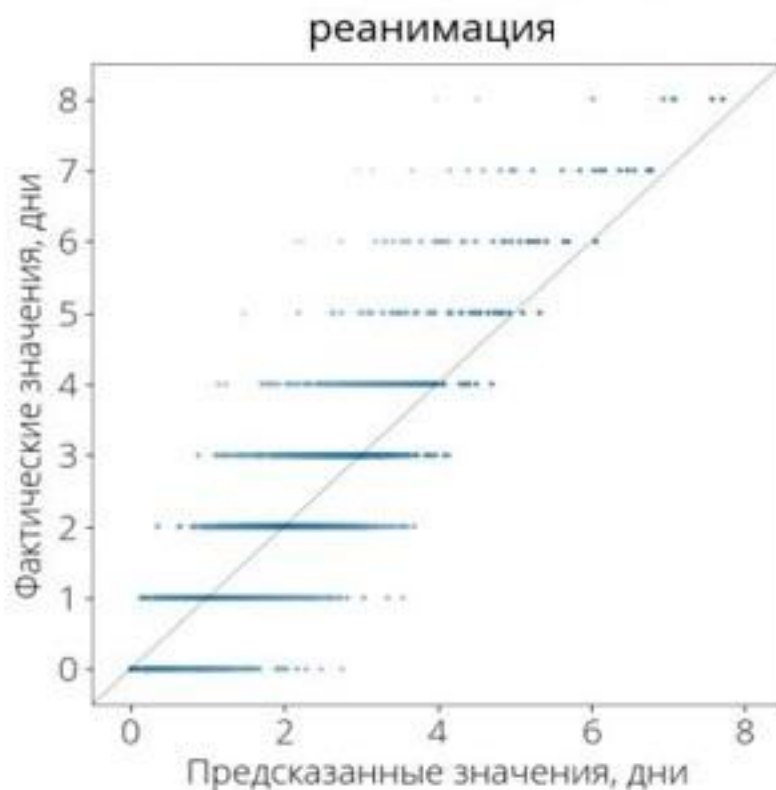
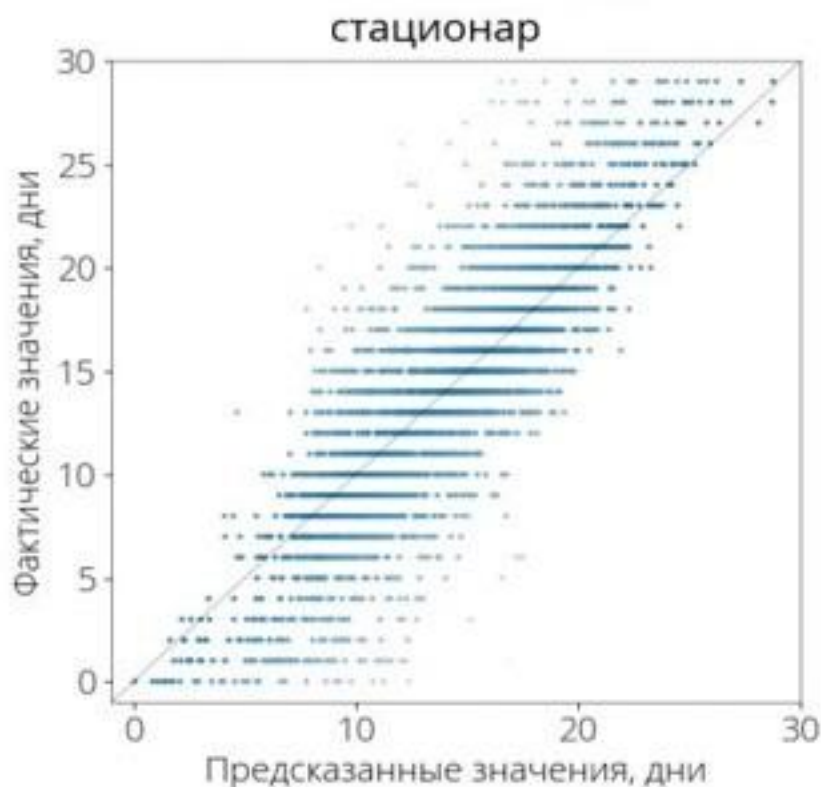
Интерпретация неоднородности: выявление связей между нозологиями по большим данным



Выявление кластеров взаимовлияющих заболеваний (пациенты с артериальной гипертензией, 89 тыс. человек): построение прогнозов о течении и развитии сопутствующих заболеваний при переходе пациента из одной возрастной группы в другую



Модельный прогноз течения эпизодов ОКС на основе нейронных сетей

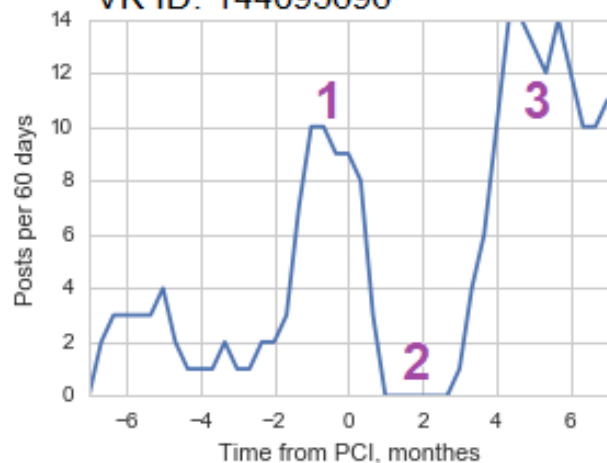


91% правильность, 92% точность, 59% полнота

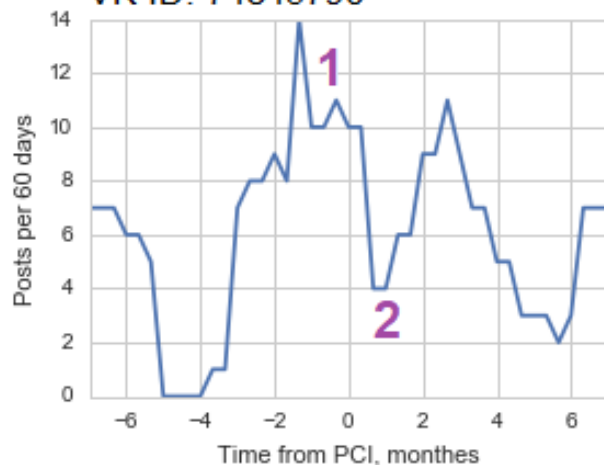
Использование данных социальных медиа: оценка влияния медицинских процедур на активность

Данные Вконтакте.RU, плановое стентирование, 543 пациента за 2015-2016 г.

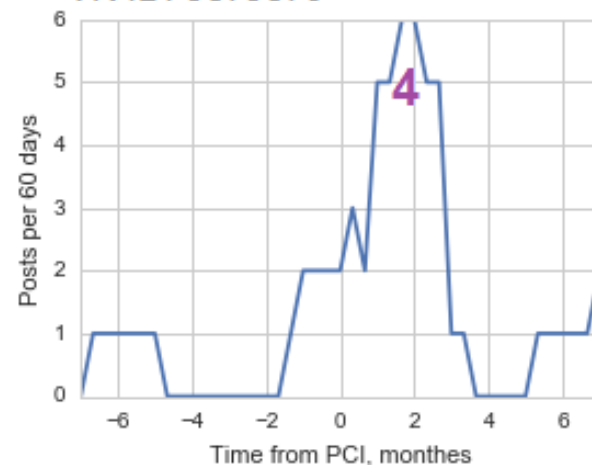
Пациент Музакка, ж, 1963 г.р.
ЧКВ 12.03.2015 г.
VK ID: 144095696



Пациент Евсенкин, м, 1953 г.р.
ЧКВ 04.03.2015 г.
VK ID: 74845790



Пациент Самигуллин, м, 1962 г.р.
ЧКВ 26.02.2015 г.
VK ID: 9879876



- (1) Относительный подъем активности в соцсети в период перед операцией (предположительно связано с низкой физической активностью)
- (2) Отсутствие активности в период непосредственно после операции (предположительно в период реабилитации)
- (3) Возможны изменения в среднем уровне активности после перерыва (в перспективе может быть интерпретировано в контексте изменения структуры активности пациента и качества его жизни)
- (4) Относительно редкий паттерн – всплеск активности после проведения операции (предположительно, активность в соцсети в ходе реабилитации (в т.ч. из стационара))

Использование данных социальных медиа: психология и психиатрия в неформальных сообществах

Психосоматические пациенты: анализ «самолечения» в
социальных сетях



Несколько
«неправильных»
слов:

- Фрейд
- Эго
- Деньги
- Интроверт
- Манипулятор
- Несчастье
- Серотонин

.....

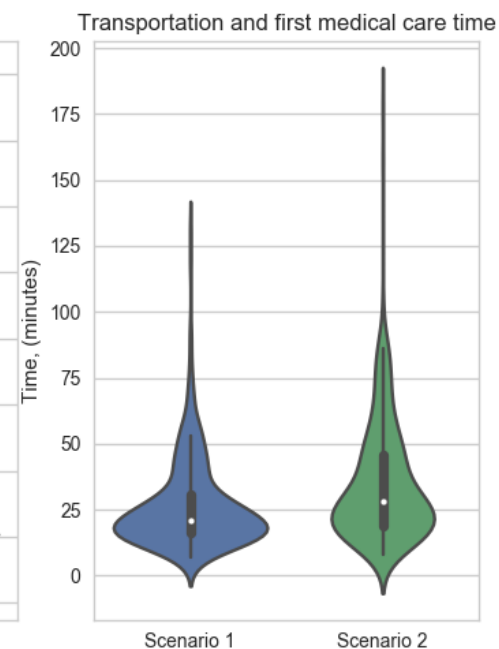
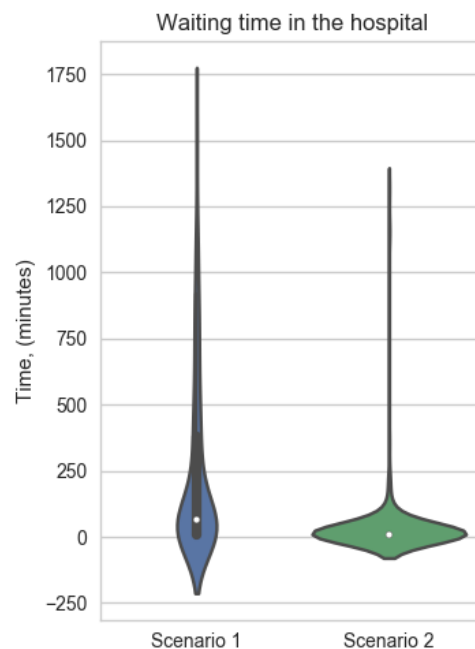
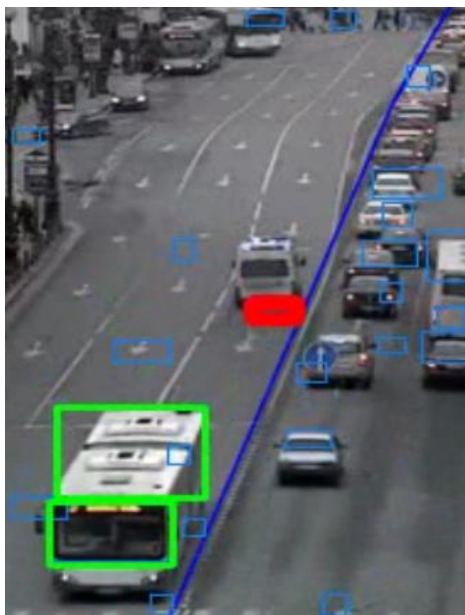
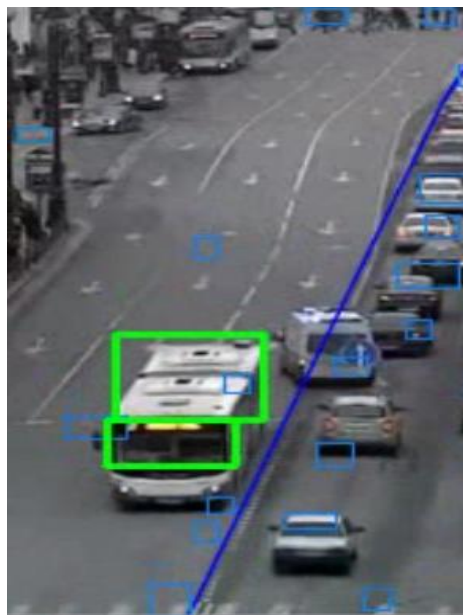
Оценка эмоционального состояния пользователей социальных сетей

ВИДЕО

(социальные сети)

Моделирование городской среды в задачах здравоохранения: флот скорой помощи

Имитационная модель диспетчеризации и маршрутизации флота скорой помощи: снижение времени ожидания в очереди на операцию на 25% за счет планирования маршрутов доставки в стационар

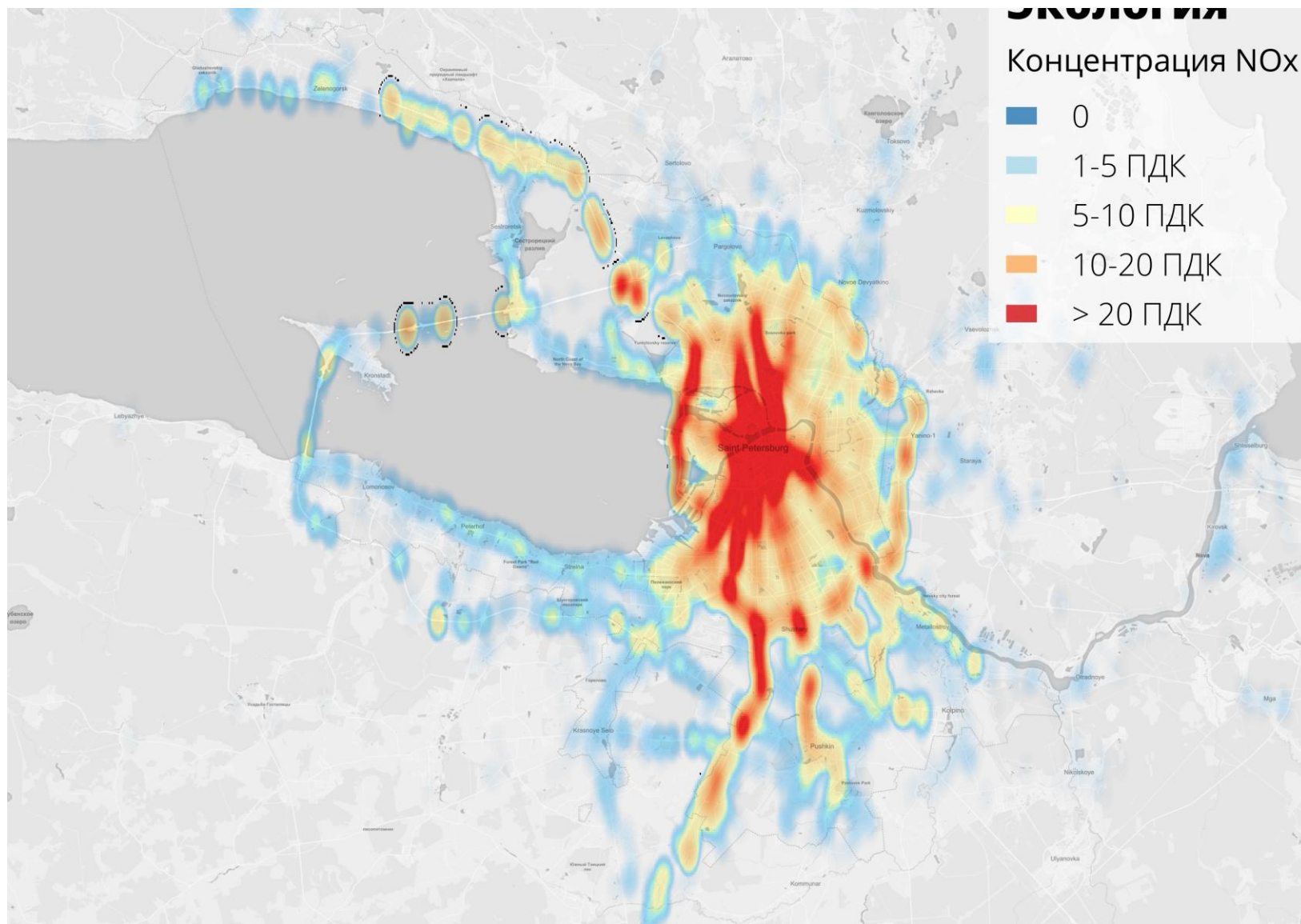




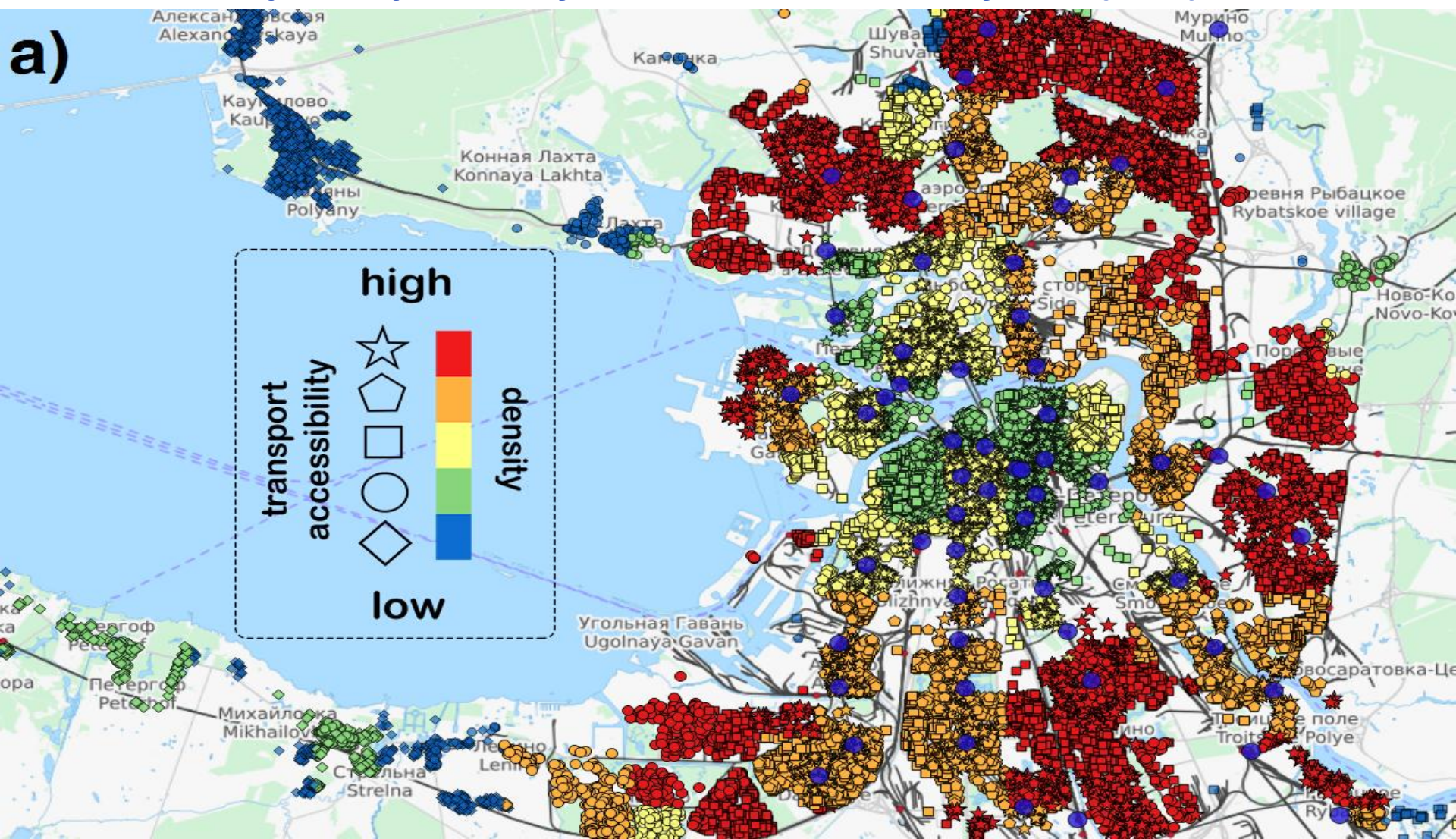
ВИДЕО

Моделирование
города «Южный»

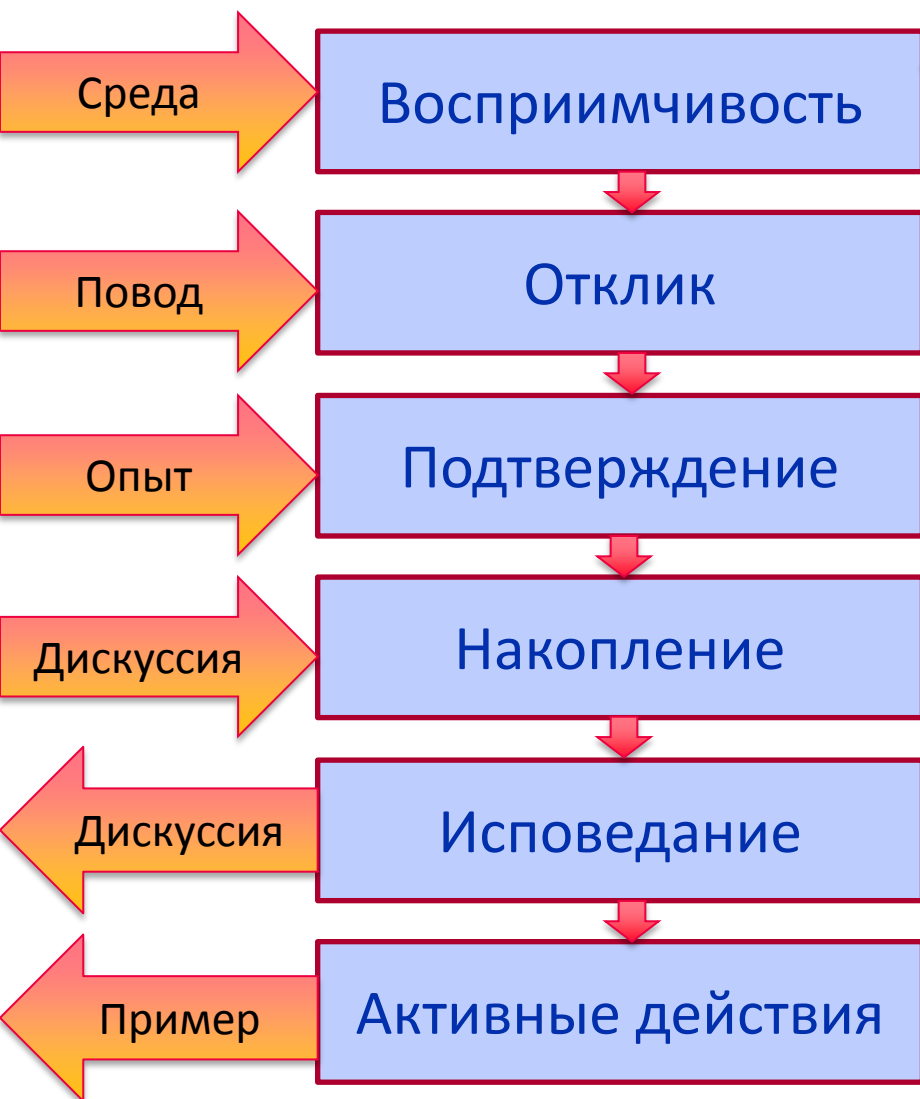
Прогнозная карта загрязнения СПб оксидами азота



Транспортная доступность: влияние социального фактора на душевное благополучие (1/2)



Транспортная доступность: влияние социального фактора на душевное благополучие (2/2)

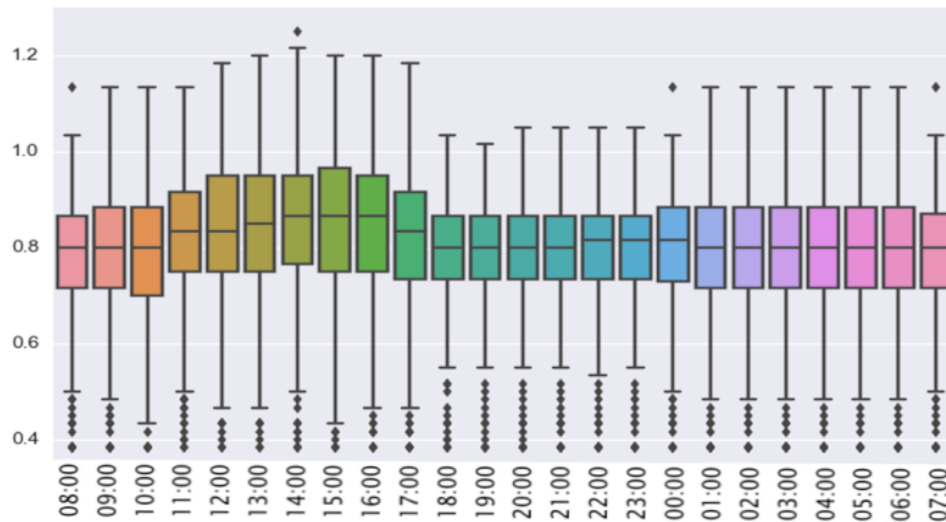
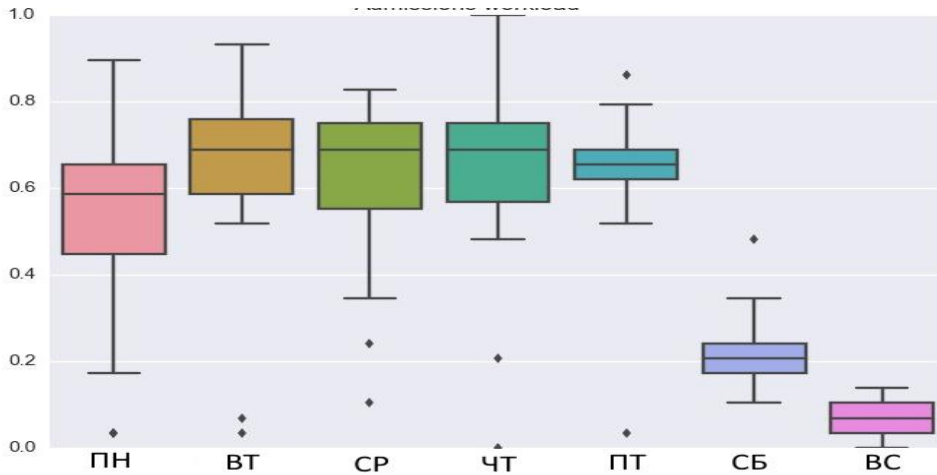


Изменение психологического состояния человека под воздействием внешней среды

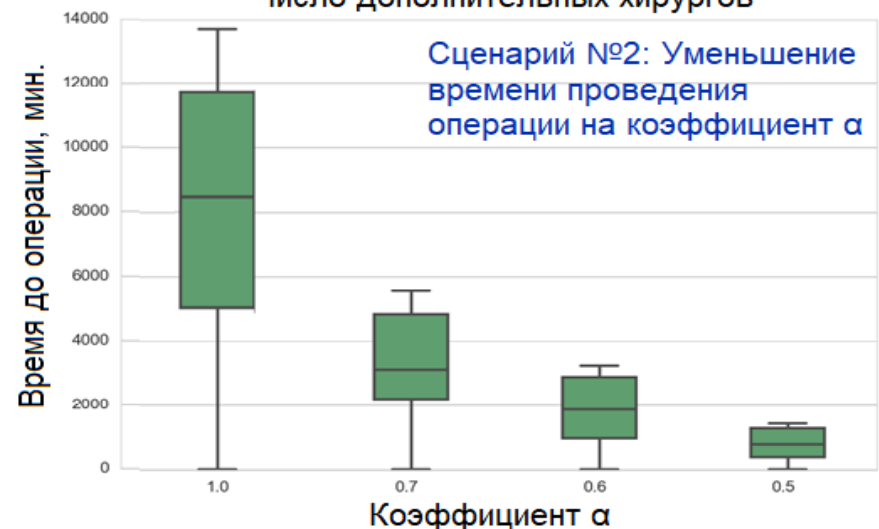
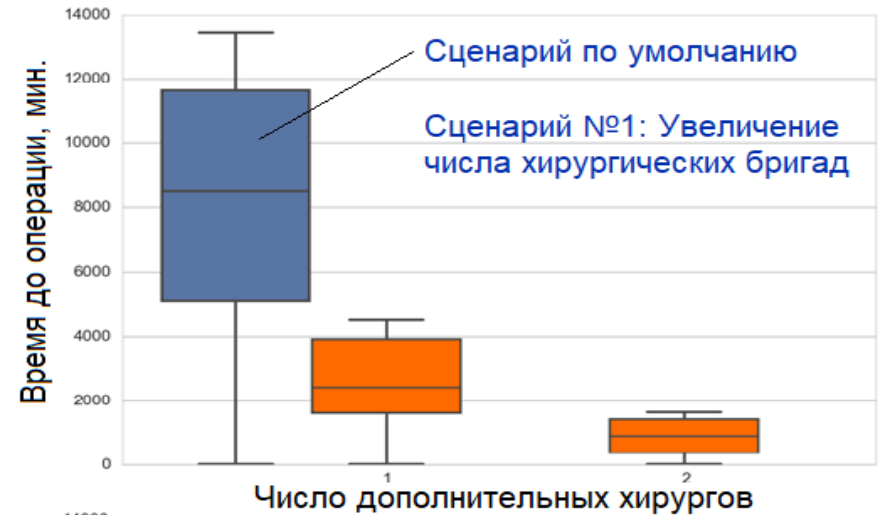


Использование предсказательных моделей для оптимизации процессов медицинского центра

Оптимизация расписаний медперсонала



Оптимизация работы хирургических бригад



Вместо заключения: почему так сложно внедрять ценностно-ориентированный подход

1. Изменение требований к архитектуре: от простых корпоративных МИС – к системам Big Data
2. Априорная «мусорность» данных: от работы с формальной статистикой – к автоматическому выделению знаний из разрозненных данных
3. Мультидисциплинарность: от изолированных регистров – к открытым пространствам и биржам данных
4. Партиципаторность: от пассивных наблюдений – к регистрации осознанных реакций пациента
5. Прогностичность: от выводов на основе только аналитики больших данных – к сценариям, созданным на основе предсказательных моделей
6. Интеллектуальность: от формального представления результатов анализа (таблицы, графики) – к СППР