

## Направление «Большие данные»

### Тема проекта № 1 «Умное общественное здание»

Современные концепты умного здания включают в себя такие технологии, как: indoor-навигации, мониторинг перемещения посетителей, распознавание посетителей, оценка состояния посетителей. В некоторых случаях также рассматривается соблюдение графика работы персонала, соблюдение рабочей формы, соблюдение мер безопасности и многое другое.

**Цель проекта** - разработать автономную платформу, реализовывающую концепцию умного общественного здания с применением технологий искусственного интеллекта.

#### Основные задачи:

1. Ознакомиться с существующими решениями, изучить источники, провести анализ ситуации. В том числе рассмотреть:
  - a. существующие системы трекинга техники по территории крупных инфраструктурных объектов;
  - b. контроль посещения зданий;
  - c. контроль средств индивидуальной защиты в общественных зданиях;
  - d. проверка униформы;
  - e. алгоритмы распознавания людей.
2. Сформулировать проблему существующих методов и инструментов.
3. Предложить решение для сформулированной проблемы. Описать основные элементы решения. Провести анализ решения, в том числе разобрать такие аспекты как:
  - a. Преимущества
  - b. Недостатки
  - c. и т.д.
4. Разработать макет предлагаемой системы.
5. Смоделировать работу системы на примерах. Провести тестовые испытания прототипа.

#### Материалы для подготовки:

1. Технологии распознавания лиц:  
<https://www.popmech.ru/technologies/402352-na-lice-napisano-kak-rabotaet-kompyuternoe-raspoznavanie-lic/>

2. Расположение систем видеонаблюдения на объекте:  
<http://videoservice66.ru/interesnoe/kak-raspolozhit-videokamery-na-obekte.html>
3. Алгоритмы машинного обучения для распознавания лиц Rosten E., Porter R., Drummond T. Faster and better: A machine learning approach to corner detection // Pattern Anal. Mach. Intell. IEEE Trans. On. 2010. Vol. 32, № 1. P. 105–119.

## Тема проекта № 2

### «Интеллектуальная система анализа настроения населения»

На сегодняшний день не существует оперативных инструментов мониторинга, которые бы наглядно показывали мнение населения и его удовлетворенность. На текущий момент социальный комфорт приходится оценивать вручную.

**Цель проекта** - создать программу автоматического анализа настроения населения.

#### Основные задачи:

1. Изучить источники, найти существующие методы и инструменты анализа настроения населения.
2. Сформулировать проблему существующих методов и инструментов.
3. Предложить конструкцию/архитектуру предлагаемого решения. Описать основные элементы решения, например:
  - a. блок определения тонов населения с помощью ключевых слов;
  - b. блок работы с текстовыми входными данными;
  - c. методы машинного обучения в области анализа текста.
4. Разработать макет предлагаемой системы.
5. Провести тестовые испытания макета.

#### Материалы для подготовки:

1. Тональный анализ:  
[https://nnov.hse.ru/data/2019/02/21/1194654193/Rybakov\\_V.pdf](https://nnov.hse.ru/data/2019/02/21/1194654193/Rybakov_V.pdf)  
<https://habr.com/ru/company/mailru/blog/516214/>  
<https://habr.com/ru/company/mailru/blog/516726/>  
<https://habr.com/ru/company/mailru/blog/516730/>
2. Методы автоматической классификации текстов:  
<https://cyberleninka.ru/article/n/metody-avtomaticheskoy-klassifikatsii-tekstov>
3. Обработка текста методами машинного обучения:  
<https://www.machinelearningmastery.ru/machine-learning-text-processing-1d5a2d638958/>

## Тема проекта № 3

### «Рекомендательная система сериалов и фильмов»

В современном мире каждый день появляется большой объем информации, в т.ч. контент для различных развлекательных медиа платформ. Для больших компаний, таких как Netflix и Youtube (Google) ежесекундно встает задача подбора контента персонально для каждого человека исходя из его интересов и других факторов. Очевидно, что предлагать новые видео или фильмы необходимо персонально каждому человеку. Предлагаем вам построить и реализовать алгоритм, который будет делать это автоматически.

**Цель проекта** - разработка автономной системы, которая будет по предыдущей активности человека предлагать ему новые фильмы, сериалы и шоу.

#### Основные задачи:

1. Ознакомиться с существующими решениями, изучить источники, провести анализ ситуации. В том числе рассмотреть:
  - a. Классические рекомендательные системы;
  - b. Алгоритмы визуализации данных.
2. Сформулировать проблему существующих методов и инструментов.
3. Предложить конструкцию/архитектуру предлагаемого решения. Описать основные элементы решения.
4. Разработать макет предлагаемой системы, в том числе визуализировать данные и реализовать прототип автономной системы.
5. Провести тестовые испытания макета системы.

#### Материалы для подготовки:

1. База данных пользователей Netflix:  
<https://www.kaggle.com/shivamb/netflix-shows>
2. Алгоритмы построения рекомендательных систем:  
<https://habr.com/ru/company/piter/blog/350346/>
3. Как работают рекомендательные системы:  
<https://habr.com/ru/company/yandex/blog/241455/>
4. Как визуализировать данные:  
<https://medium.com/nuances-of-programming/5-%D0%BF%D1%80%D0%BE%D1%81%D1%82%D1%8B%D1%85-%D1%81%D0%BF%D0%BE%D1%81%D0%BE%D0%B1%D0%BE%D0%B2->

[%D0%B2%D0%B8%D0%B7%D1%83%D0%B0%D0%BB%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D0%B8-%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85-%D0%BD%D0%B0-python-%D1%81-%D0%BA%D0%BE%D0%B4%D0%BE%D0%BC-e0053808c83d](#)