

Папамс

кейс 3

лучшая концепция
лучшее решение на code

Басалаев
Максим
87712151111
[tg@pojaleesh](tg://user?id=pojaleesh)

Партнеры

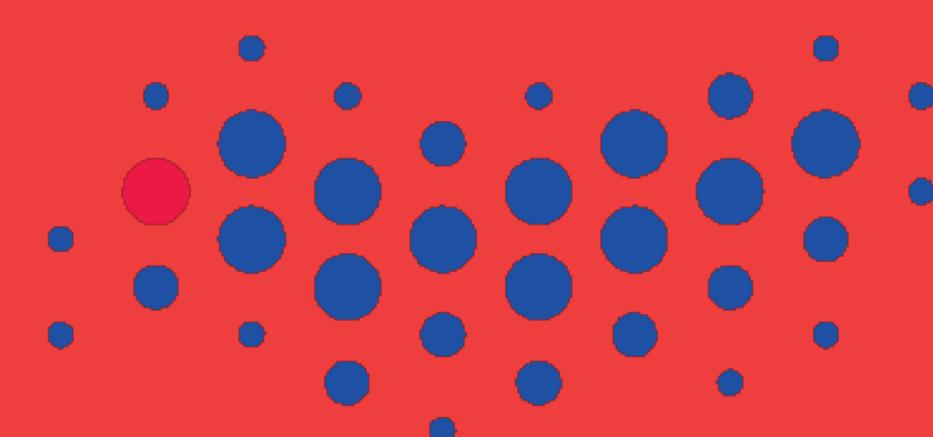


ЦЕНТР
ТРУДОВЫХ
РЕСУРСОВ



Правительство Санкт-Петербурга
Комитет по труду
и занятости населения
Санкт-Петербурга

ТВОЙ **ПЕРВЫЙ**
ШАГ В КАРЬЕРЕ >>>



УНИВЕРСИТЕТ ИТМО



OXYGEN



парсинг

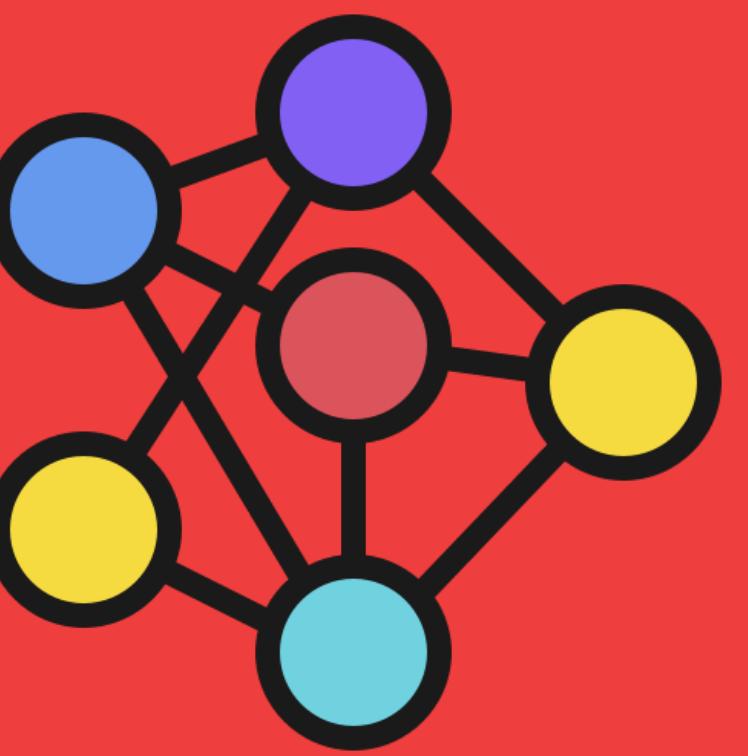


парсинг



парсинг

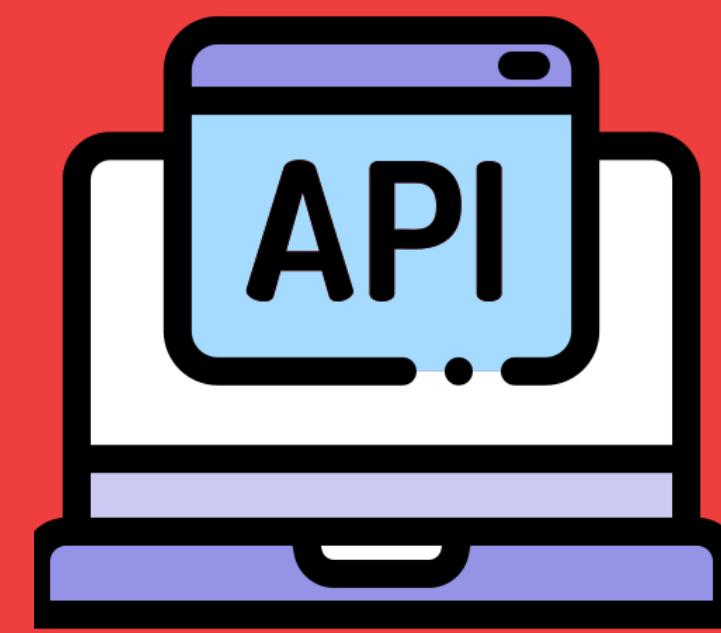
Используем алгоритм на сервере



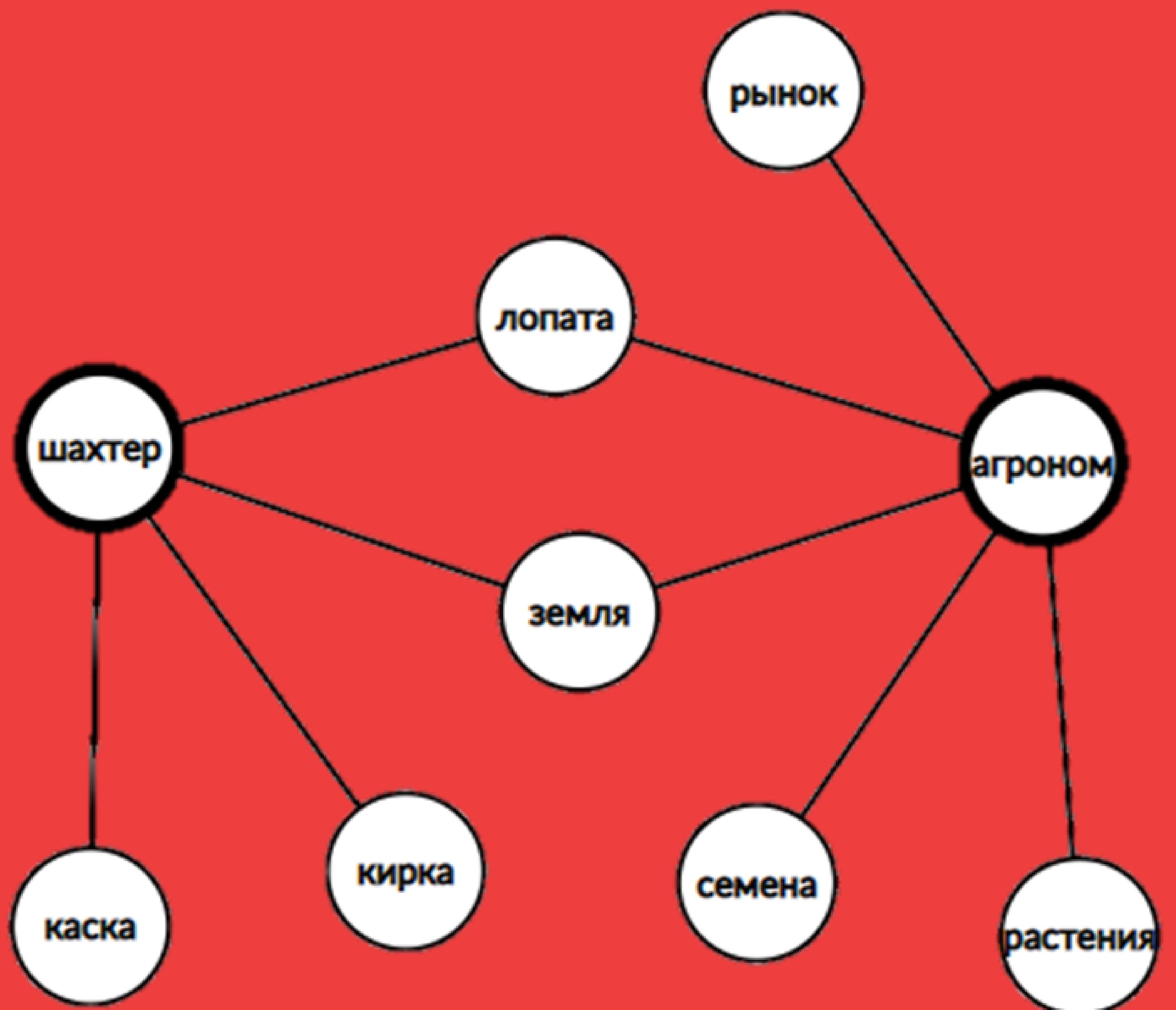
Получаем профессию
человека и его навыки

Парсинг

По возможности мы использовали арі, но в некоторых случаях было удобней использовать библиотеку selenium



Алгоритм



Используя парсеры, получили большую выборку профилей с конкретной профессией для обучения

В итоге для каждой профессии получили ключевые слова, которые ее характеризуют

Построили граф, в котором веса - это суммарное число вхождений ключевого слова для каждой профессии

Алгоритм

Теперь для определения профессии нового профиля, для каждого слова из профиля ищем похожие по смыслу слова в графе подобия слов. Мы использовали word2vec из библиотеки gensim

Для быстрого обнаружения слова мы использовали алгоритм Ахо-Корасик, затем если оно было найдено, то мы прибавляли к профессии соответствующий вес с помощью метода tf-idf

Что дальше?



Посредством сотрудничества с социальными сетями, можно получать больше данных



Получение более точного алгоритма



→ Возможность быстрого поиска штата
сотрудников, а также достижимость
монетизации этого дела





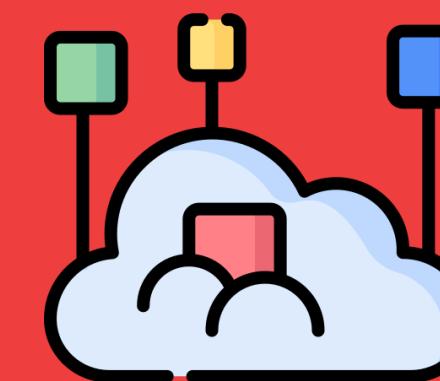
Какие же нужны
ресурсы?



Финансовые



Больше данных!

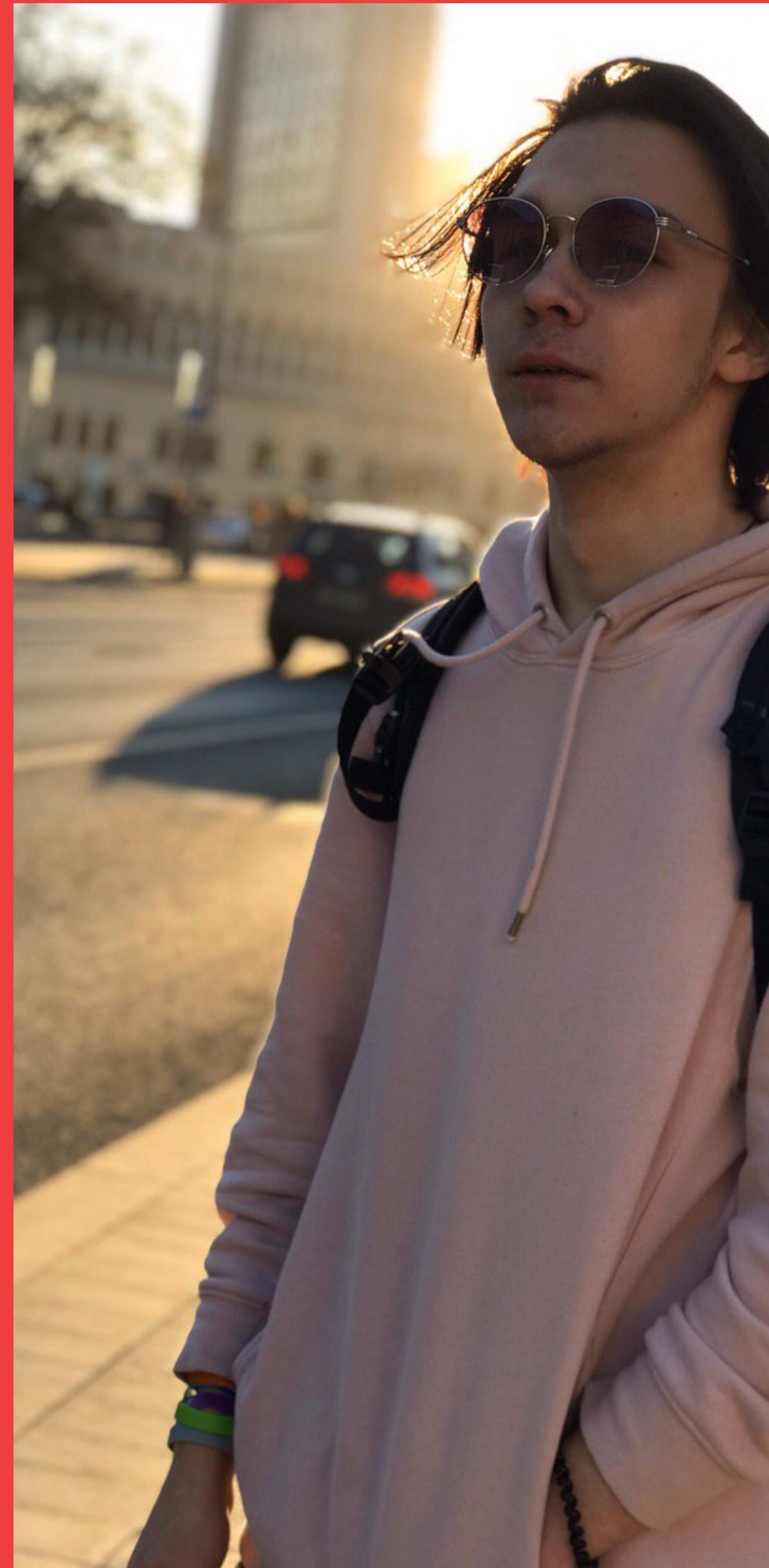


Папамс



Артем Мельников

data-science



Иван Лущ

software developer



Максим Басалаев

algorithm and math



Адильхан Кусиденов

software developer

Учимся в НИУ “Высшая школа экономики” на программе прикладная математика и информатика.



Спасибо за внимание!