

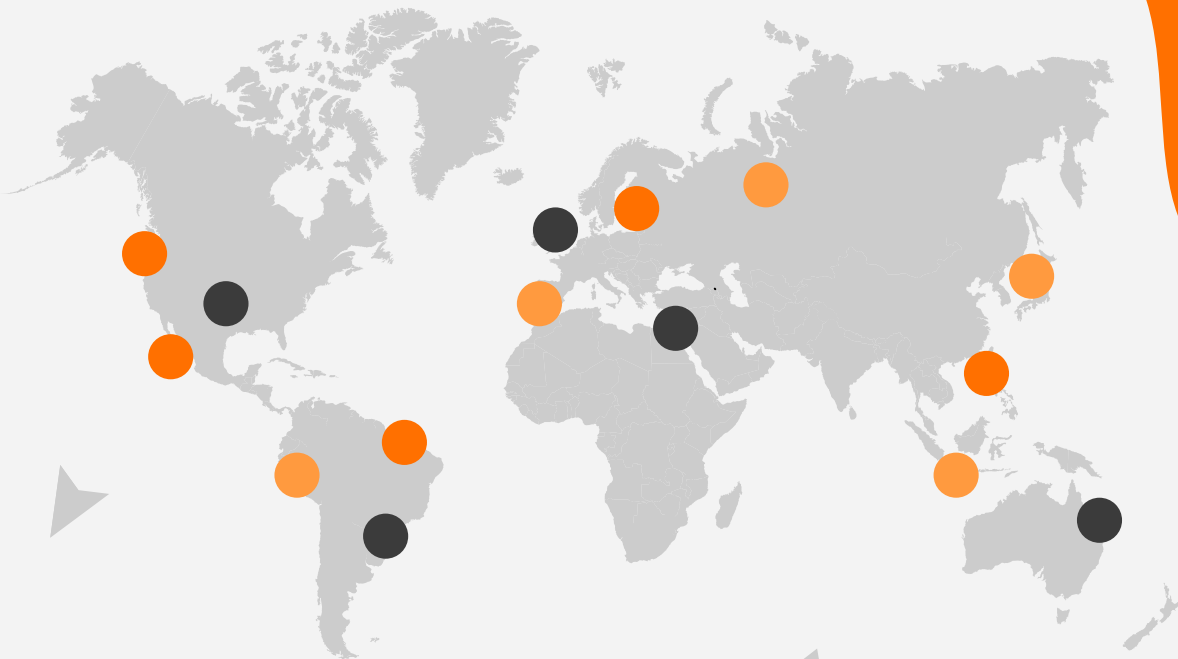
# Определение корректного адреса в базе данных

Команда **Punk Butterfly**

# Почему задача важна



1. Улучшение функциональности существующих и реализующихся сервисов;
2. Коррекция и унификация множества используемых городских баз данных.



# На что делали упор

**01**

## Точность

Поиск наиболее соответствующих адресов

**02**

## Масштабируемость

Простая адаптация для новых адресов

**03**

## Актуальность

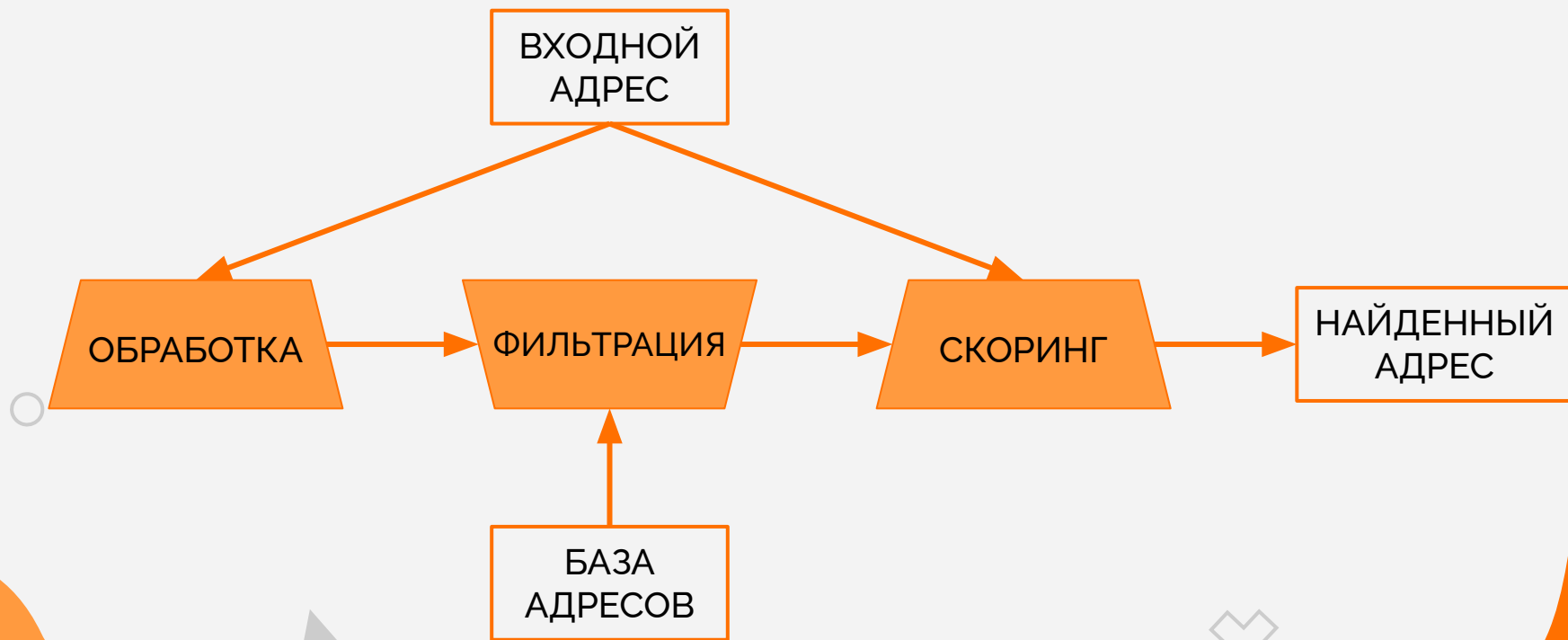
Современный стек технологий и подходов

**04**

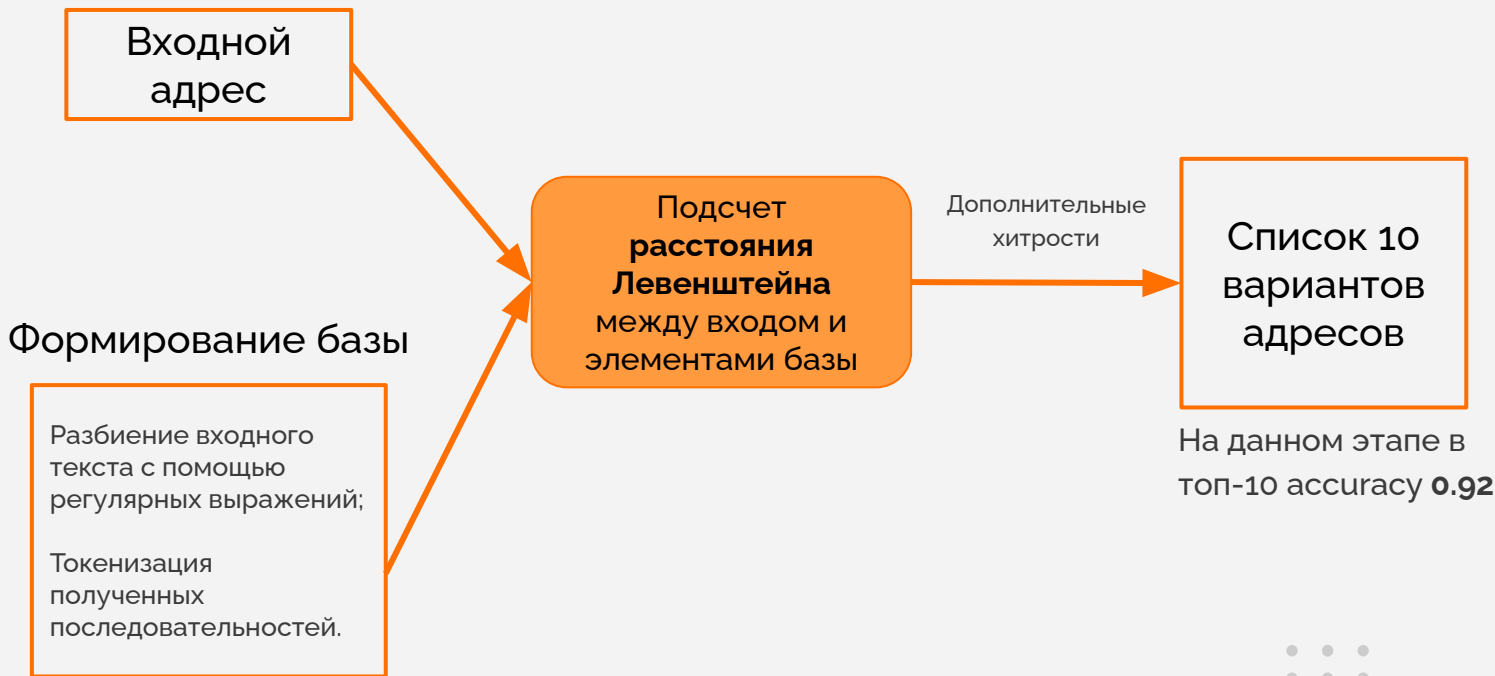
## Гибкость

Контроль составных частей архитектуры

# Архитектура модели



# Обработка и фильтрация



# Генерация датасета v1

город Петергоф,  
Собственный проспект,  
дом 64, строение 1

ИСХОДНЫЙ  
АДРЕС

town\_name

town\_shortcode

town\_onlyname

liter\_name

liter\_shortcode

liter\_onlyname

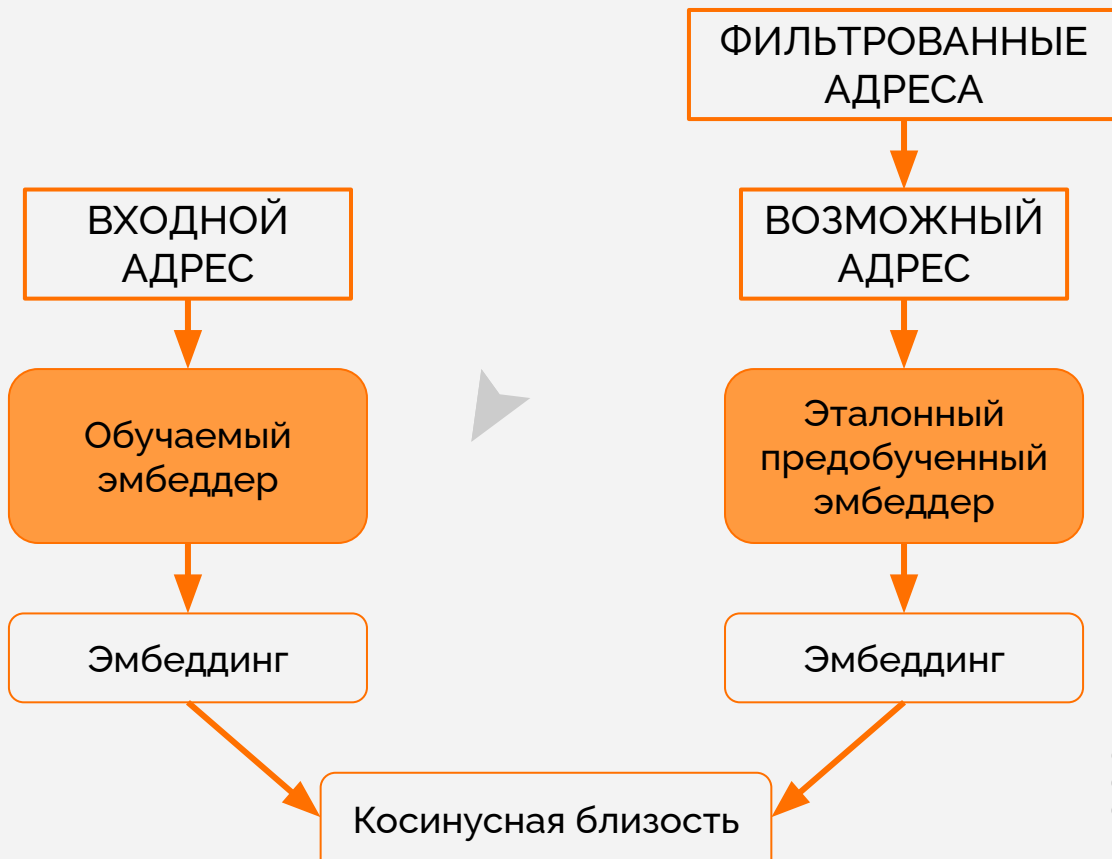
Вероятностный подход

СОЗДАНЫЙ  
АДРЕС

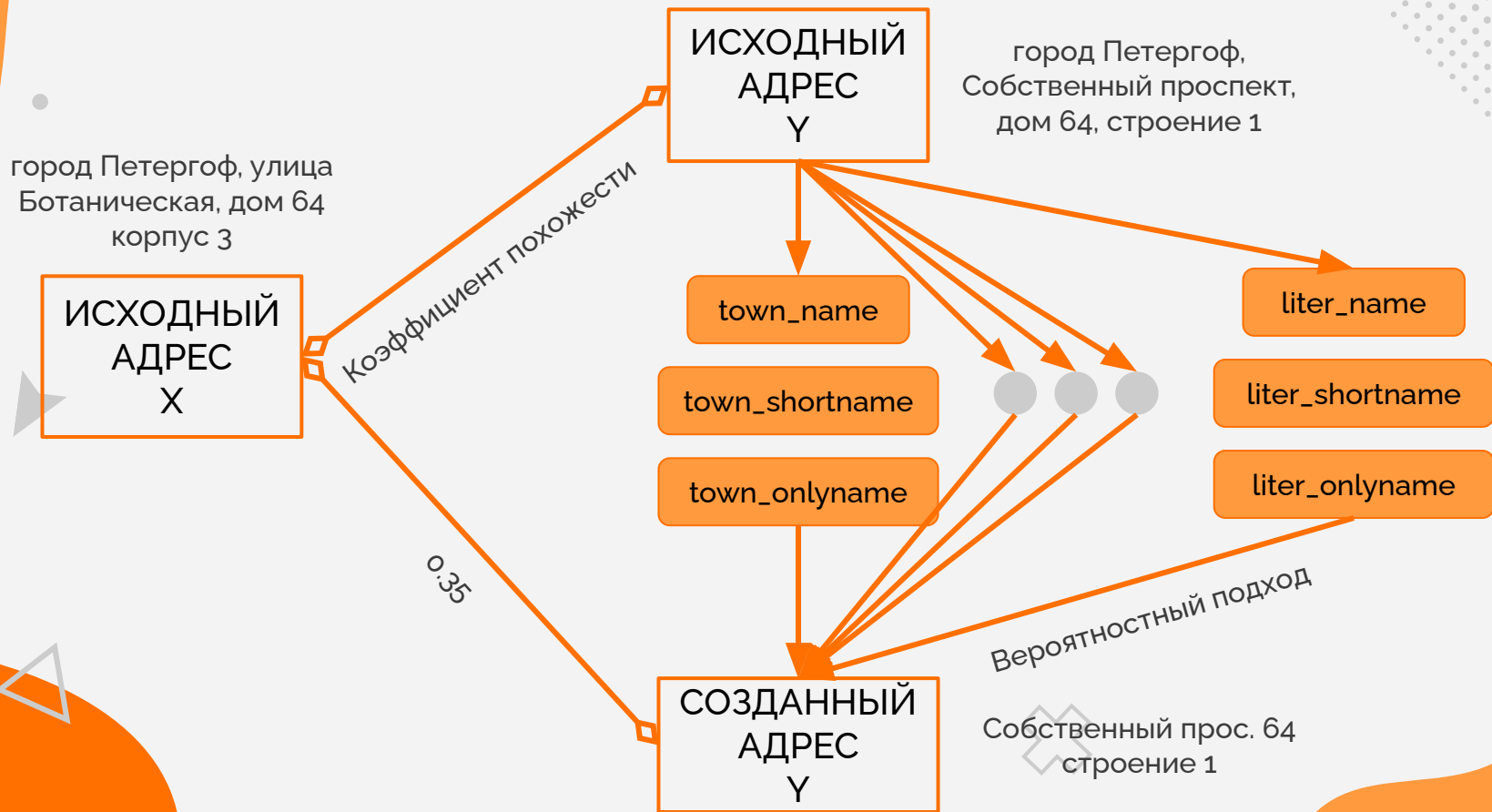
Собственный прос. 64  
строение 1



# Скоринг v1

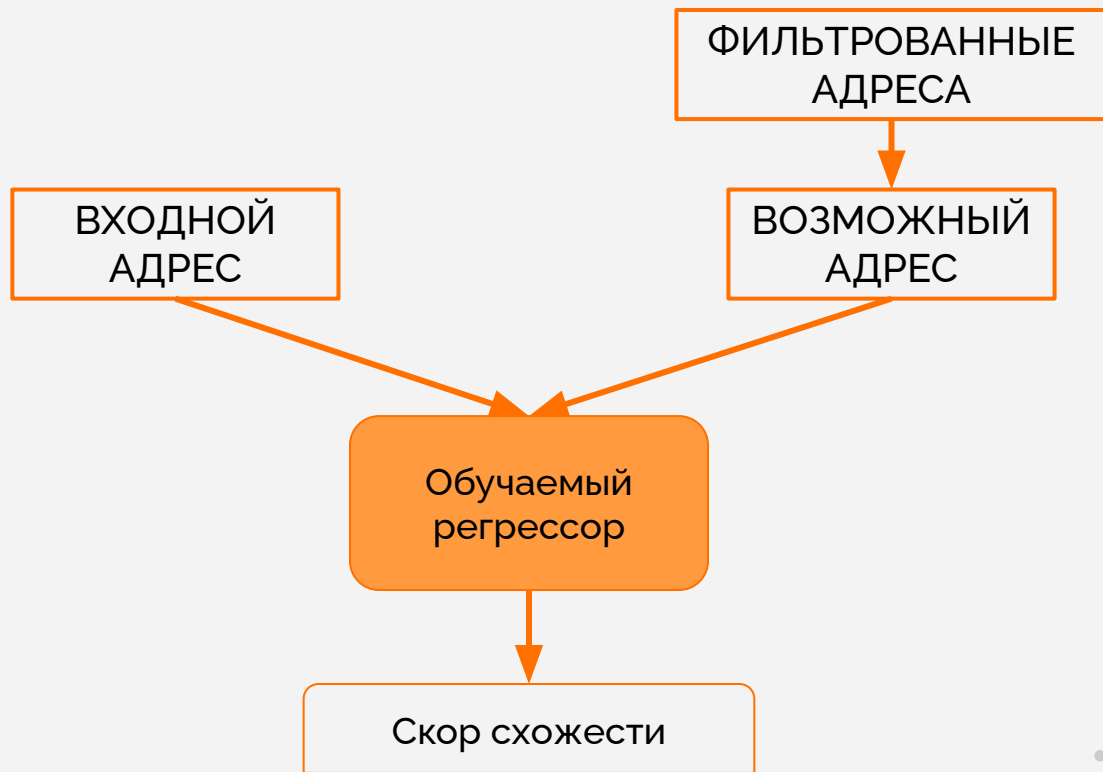


# Генерация датасета v2





# Скоринг v2



# Точность подходов

**72%**

**TOP 1**

**92%**

**TOP 10**

# Преимущества решения

Дообучение

Транслитерация

Интерфейсы



МАСШТАБИРУЕМОСТЬ



# Технологии

Fast API

Streamlit

Pytorch

Docker

Transformers

Yandex  
Cloud

# Команда Punk Butterfly

**Роман  
Исаев**

Аналитик - джойнер

**Илья  
Куприянов**

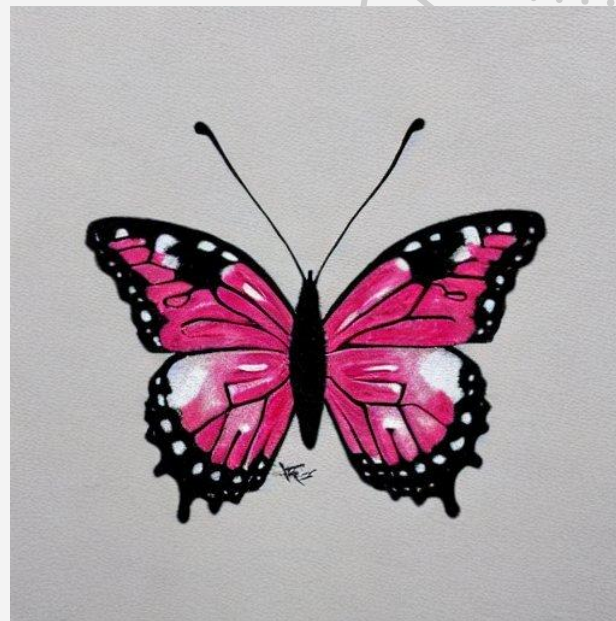
Докер - магистр

**Алексей  
Полещенков**

MVP - алгоритмист

**Александр  
Школин**

Архитектор - теоретик



# Контакты



WEB - сервис



TG Команды