

IMPRECITY

4SPB

Hackathon

16-17 февраля 2019

CONCEPT 1

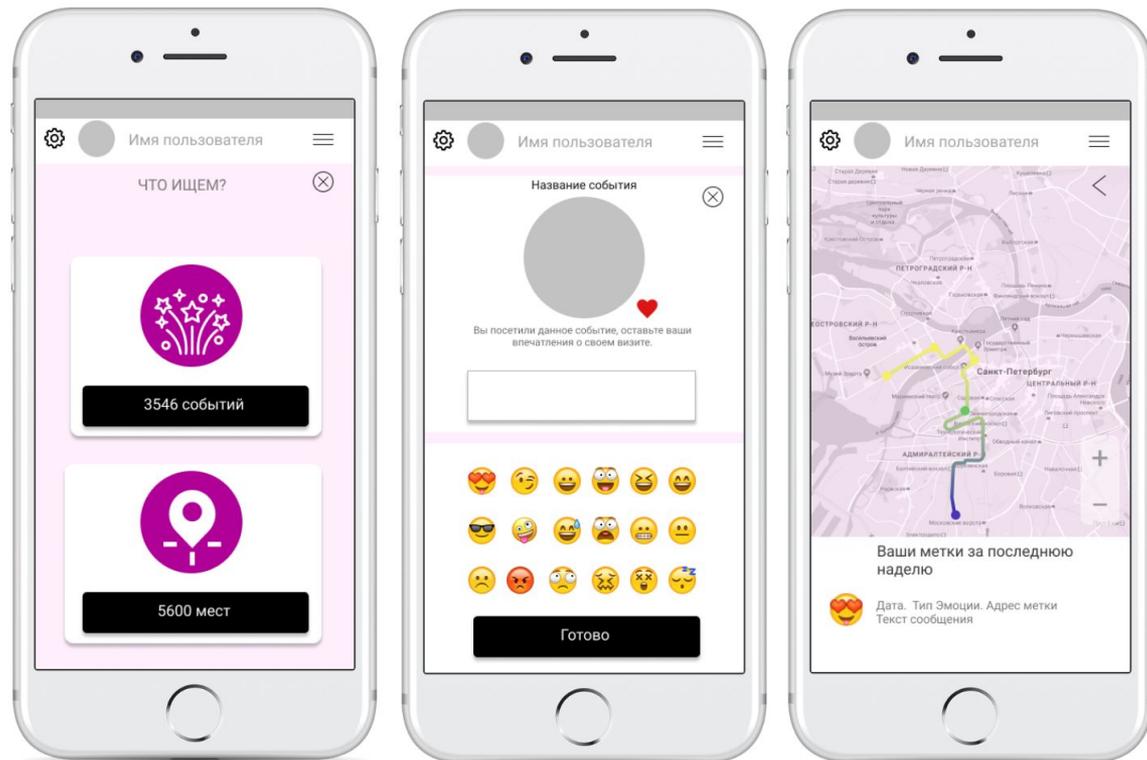
Для изучения эмоций города мы разработали концепт приложения **Imprecity**, которое на данный момент запущено в тестовой веб-версии. **Imprecity** - это персональный дневник, где пользователь регулярно отмечает свои эмоции и впечатления об общественных пространствах в городе.

В основу заложена концепция базовых эмоций Пола Экмана - радость, злость, грусть, страх, отвращение и удивление.

Приложение выдает статистику и эмоциональные треки пользователя за месяц. Это позволяет установить взаимосвязь между опытом пользования городской средой и субъективным состоянием человека.

Больше информации о проекте <https://imprecity.ru> <http://imprecity.tilda.ws/>
<http://news.ifmo.ru/ru/news/7538/>

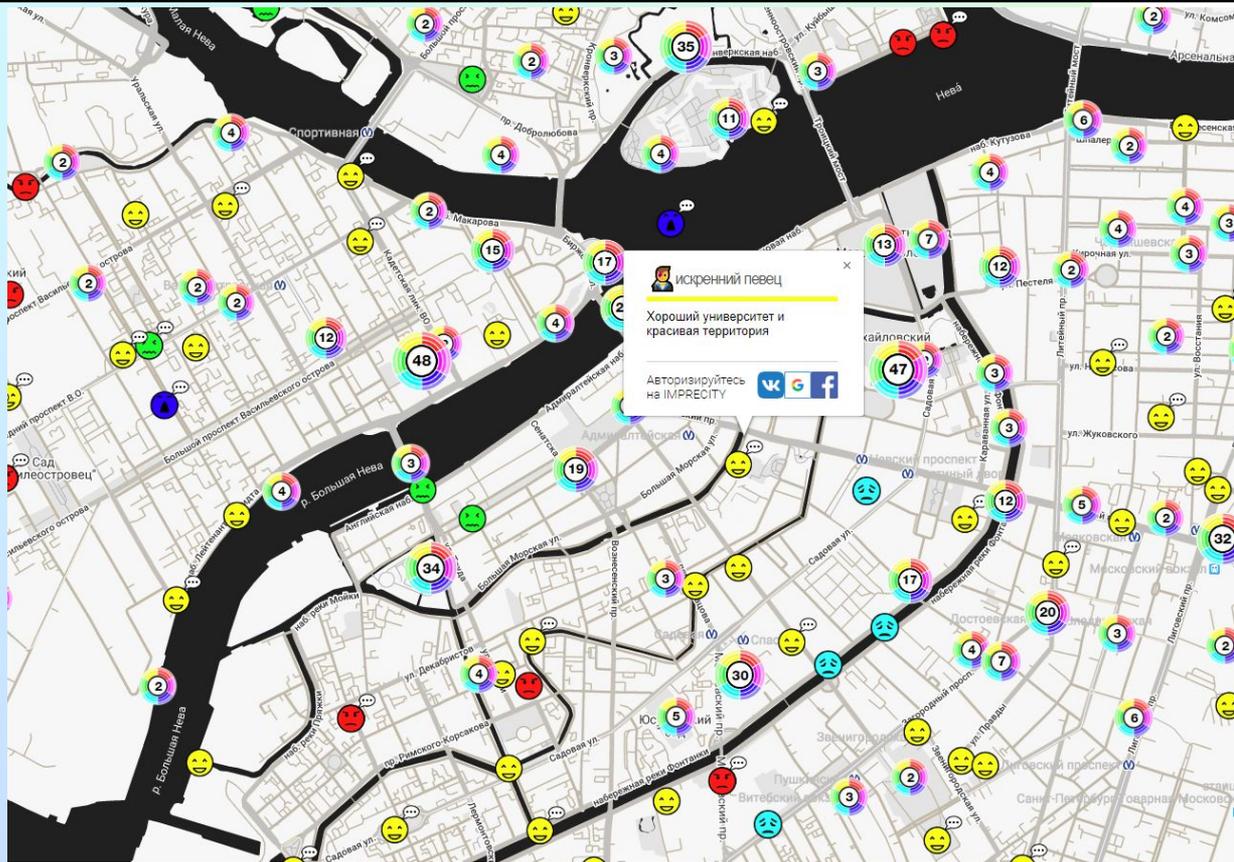
CONCEPT 2



Прототип приложения

WEB

На текущий
момент в web
версии
Imprecity
собрано **1843**
эмоциональн
ых метки и
1018
комментариев
к ним.



TASK



- Команда работает над алгоритмом автоматизированного распознавания эмоций **6** классов - радость, грусть, злость, отвращение, страх, удивление, в коротких текстах пользовательских отзывов.
- Основа “истинных данных” для алгоритма - отзывы посетителей **Imprecity (1018)** и размеченная экспертами выборка отзывов **Google Places (2450)**.
- Планируется применить алгоритм к данным **Tripadvisor (более 50К записей)**.

ALGORITHM 1.0.

На основе обучающей выборки **Imprecity** и **Google Places** был обучен алгоритм наивного Байеса. Точность алгоритма была проверена на контрольной выборке **Imprecity** и **Google Places**.

DATA 1

Размеченные данные по **6** классам эмоций:

Imprecity - **1018** отзыва

Google Places - **2450** размеченных экспертами отзыва

Данные для дальнейшей разметки:

TripAdvisor - **50767** комментариев **1290** мест Санкт-Петербурга и Ленинградской области взятых с **tripadvisor**. Комментарии оставлены в период с **2005-08-24** по **2019-02-05**

DATA 2

Структура **json** файла с данными

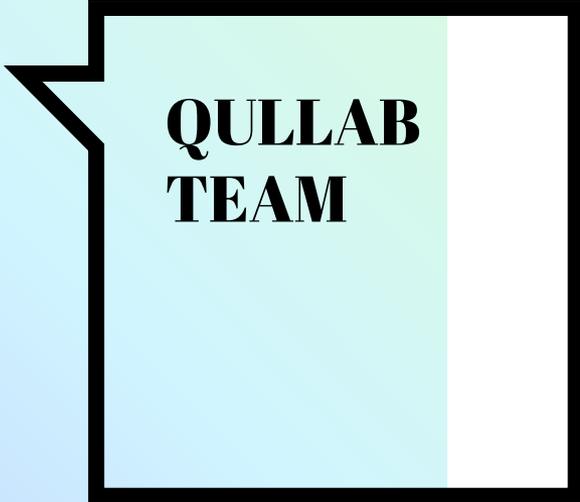
```
{'name','rating','lat','lng',  
'reviews': [{'date','rating','text',  
              'mark' // номер эмоции  
              },...]  
},...]
```

PROBLEM



При проверке с помощью матрицы неточностей наивный байесовский алгоритм показывает точность **80%** на **2** классах эмоций - радость и остальные негативные эмоции, **НО** дает ниже **50%** точности на **6** классах эмоций.

Необходимо оптимизировать алгоритм распознавания эмоций по **6 классам в текстах пользовательских отзывов. На Хакатоне мы хотим попробовать решить эту задачу вместе с вами**



**QULLAB
TEAM**

Александра Ненько

социолог, урбанист

Артем Конюхов

программист, аналитик
городских данных

Петр Ревякин

IT специалист,
разработчик

Марина Петрова

аналитик городских
данных, дизайнер

QULLAB

Лаборатория качества
городской жизни ИТМО

www.qullab.com

www.imprecity.ru

