Cassandra. Front-end.

# Варианты запросов в систему

## Лояльность

1. Отношение к субъекту
2. Отношение к субъектам (например, к главам администраций, или к главам районов)
3. Отношение к событию произошедшему или к предстоящему (к инициативе)

## События

1. Распределение событий по местоположению
2. Распределение событий по категориям
3. Возрастные категории субъектов тех или иных событий (например, криминальных происшествий, культурных мероприятий и т.п.)

## Претензии

1. Претензии к конкретному субъекту (перечень, доля)

## Общая обстановка

1. Доля плохих/хороших новостей
2. Доли новостей по категориям

## Настроение жителей

1. В градации от 1 до 5

## Аналитика

1. Корреляция выборки событий с настроением жителей

# Взаимодействие пользователя с системой

* Каким образом (как и кто) вводит запросы в систему ?
* Формат запросов ? на естественном языке ? или через заполнение специализированных полей (напр. Субъект, место, событие и т.п.) ?
* Требуется ли участие специалиста, эксперта по системе, который будет вводить данные запросы ?
* Между запросом в систему и выдачей результатов ожидается период анализа, подготовки данных. Насколько длительным он может быть ?
* Отбор по категориям (?)

Один из вариантов работы – специалист, зная конечные запросы пользователя, настраивает в каком разрезе требуется обработать и проанализировать первичные данные. Данные обрабатываются, видимо достаточно длительное время. А на финише пользователь может выбрать один или несколько запросов, среди предварительно заданных, и выбрать вариант визуализации. Этот этап должен выполняться быстро.

*//в ArcGIS – результаты работы автоматически публикуются как WebService*

# Источники данных

Соц. сети – vk, fb, ok, Instagram, Twitter

Мессенджеры – tg, viber (мало открытых групп), whatsapp (мало открытых групп)

Глобальные новостные сайты (включая обсуждения) – ria, rbc, … см. по рейтингу сайтов (если есть поиск – выборка новостей по кодовому слову (напр. Челябинск, или по тэгам), если поиска нет, то перебор новостей и фильтрация

Местные новостные сайты (включая обсуждения) – 74.ru, 1obl.ru, chel.kp.ru, …

Иностранные новостные сайты – см. по рейтингу (+ требуется автоматический перевод)

Ю-туб – названия роликов, обсуждения в комментариях

# Привязка исходных данных к местоположению

Привязка к местоположению по «месту жительства» авторов сообщений затруднена, т.к. эта информация отсутствует. Чтобы ее получить требуется анализировать профиль автора, но полнота и качество такой информации предположительно будет низким.

При анализе данных в привязке к конкретному субъекту или событию возможна привязка к местоположению через сам объект анализа.

*Много обсуждений, привязанных к районам города, и/или к конкретным заведениям происходит в мессенджерах Viber, WhatsApp, в закрытых группах. В некотором смысле это кладезь информации, но доступ к ним без включения в группу невозможен.*

# Визуализация

## Визуализация исходных данных

Может потребоваться для демонстрации количественного объема выборки, распределенного по периодам времени, по местоположению.

Визуализация по местоположению может быть продемонстрирована с привязкой к карте: в центре районов выводятся точки разной толщины в зависимости от количества выборки по району, в подписи выводится само количество.

Визуализация по периодам может быть выведена в виде гистограммы.

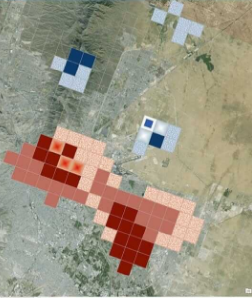
## Визуализация промежуточных данных

?

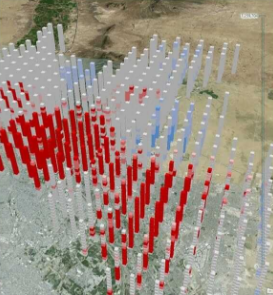
## Визуализация результатов

1. Отношение к субъекту

* При возможности привязки к местоположению
  + Вариант 1. На карте - градиентной заливкой с переходом цвета. Диапазон - максимум негатива – насыщенно красный, максимум позитива – насыщенно зеленый. Данный вариант не дает понимания количества данных первичной выборки, повлиявших на оценку.



* + Вариант 2. Карта изображается в 3D, вся карта разбивается на участки и каждый участок содержит гистограмму. Гистограммы на участках разных районов будут иметь высоту, соответствующую доле сообщений в общей выборке. Все гистограммы будут раскрашены градиентной заливкой, аналогично варианту 1.



* При невозможности привязки к местоположению. Здесь необходимо привлекать набор типовых средств и кастомизировать их к поставленной задаче, с целью предоставить возможность заказчику проанализировать текущее состояние и тенденцию. Используются круговые диаграммы, точечные диаграммы, графики, гистограммы в различном сочетании.

## Расшифровка визуализированных данных по клику

Нужен переход к первичным данным (в обработанном виде), желательно чтобы можно было бы открыть сообщение в источнике. Если первичных данных много, то м.б. сгруппировать их либо по периодам, либо источнику, либо какому-либо другому критерию.