

# Соц граф чиновника

*(telegram бот для составления списка профилей ВК  
из ближайшего круга чиновника)*

# Какую проблему решает

Во время антикоррупционных расследований часто возникает задача поиска социальных связей чиновника или, например, участников торгов на госзакупках.

Сейчас это сводится к ручному прокликиванию профилей и групп в соцсетях (что долго и муторно) а также к ручному определению релевантных для расследования профилей.

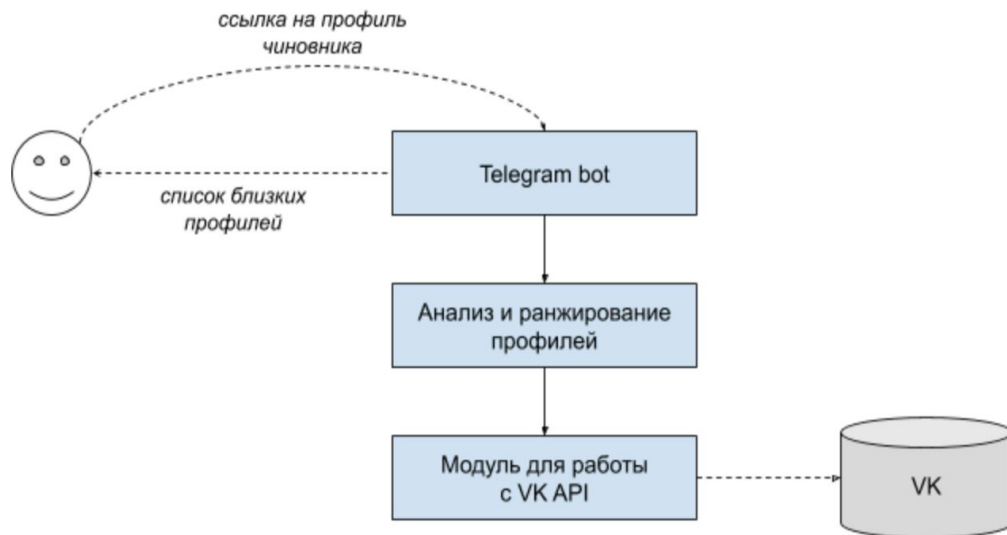
Этот проект – попытка автоматизировать работу с соц сетью ВК и написать расширяемую инфраструктуру для всего остального.

# Технические детали

## Признаки близости профилей:

- Находится ли в списке друзей
- Есть ли совпадение по фамилии (учитываются обе возможные версии)
- Совпадает ли девичья фамилия
- Указан ли в разделе с семейным положением и родственниками
- Есть ли совпадение по городу
- Есть ли совпадение по образованию
- Есть ли совпадение по военной службе
- Есть ли пересечения по подпискам
- Есть ли пересечения по пабликам

Для подсчета рейтинга профилей мы использовали количество совпавших признаков. И никакого ML.



- Принимаем на вход ссылку на страницу ВК
- Через VK API собираем открытые данные о чиновнике и о друзьях
- Вычисляем рейтинг для каждого профиля по "признакам близости"
- Ранжируем профили по рейтингу
- Отвечаем в чате отсортированным списком ближайших профилей

**Демо**

# Дальнейшие улучшения

- Второй раунд фидбэка от исследователей
- Расширить набор признаков, добавить веса в формулу подсчета рейтинга
- Добавить поддержку других соц сетей – **OK**, Facebook, **Instagram**
- Сопоставление информации из разных соц сетей
- Минималистичный сайт с **визуализацией** социального графа чиновника и списка родственных профилей с указанием совпавших признаков близости
- Собрать размеченные данные от реальных людей, предлагая из двух близких профилей выбрать тот что ближе
- ... ваши предложения? ...