

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.

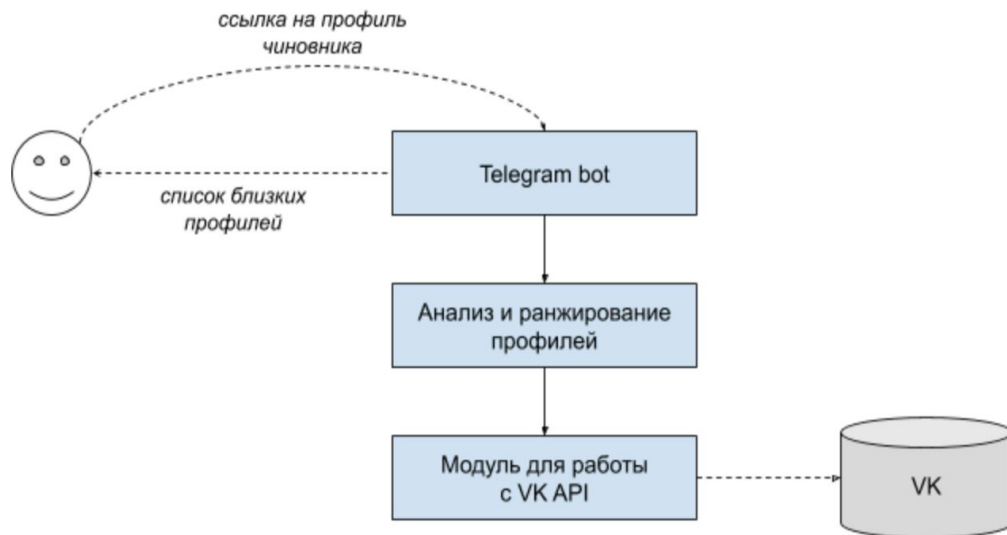
Этот проект – попытка автоматизировать работу с соц сетью ВК и написать расширяемую инфраструктуру для всего остального.

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


https://vk.com/ilya_yashin


vk
Ilya Yashin
Красносельский район Москвы.
ЦАО



7:12 AM ✓✓

 **Алексей Фролов (Челябинск)**
https://vk.com/ilya_yashin

 Принято! Начинаю анализ друзей...
Это может занять несколько минут 7:13 AM

 **Алексей Фролов (Челябинск)**
https://vk.com/ilya_yashin

✓ Готово!

1. <http://vk.com/id3420> – Георгий Лобушкин
Рейтинг: 2.0
Совпадения: друзья, образование
2. <http://vk.com/id10388> – Вера Циколия
Рейтинг: 1.0
Совпадения: друзья
3. <http://vk.com/id18698> – Константин Янкаускас
Рейтинг: 3.0
Совпадения: друзья, город, образование
4. <http://vk.com/id20952> – Кристина Констанденкова
Рейтинг: 2.0
Совпадения: друзья, образование
5. <http://vk.com/id26678> – Александр Шуршев
Рейтинг: 1.0
Совпадения: друзья
6. <http://vk.com/id106304> – Игорь Ермоленко
Рейтинг: 1.0
Совпадения: друзья
7. <http://vk.com/id114768> – Мария Преображенская
Рейтинг: 1.0
Совпадения: друзья
8. <http://vk.com/id131721> – Дуся Соболев
Рейтинг: 2.0
Совпадения: друзья, образование
9. <http://vk.com/id204876> – Маргарита Лошак
Рейтинг: 1.0
Совпадения: друзья
10. <http://vk.com/id209906> – Елена Солнечная
Рейтинг: 1.0
Совпадения: друзья

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? ...