

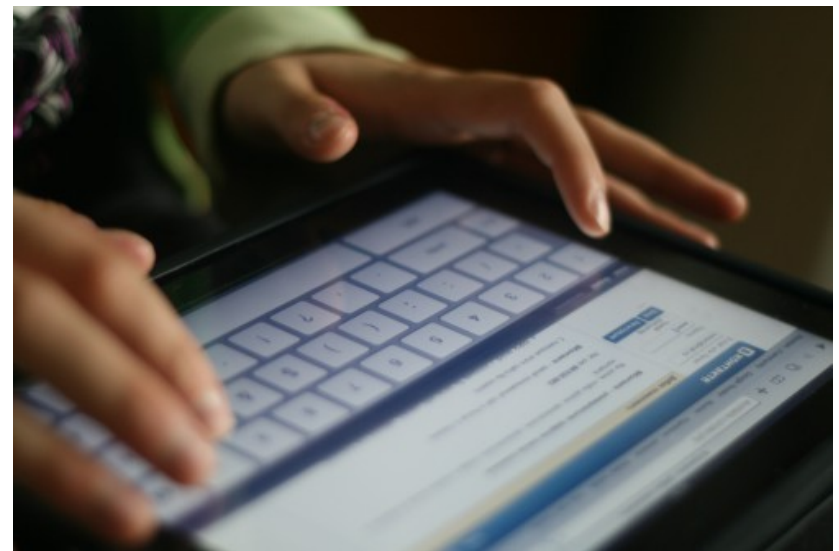
XI международная научно-практическая конференция «Электронные средства и системы управления» ТУСУР

Анализ тональности текстовых сообщений. Система мониторинга настроения людей в социальных сетях

Мейта Марина, Шиповской Владислав
Студенты кафедры КИБЭВС, ФБ ТУСУР
Руководитель: Романов А.С.

Назначение системы и ее актуальность

- **Социальные сети**
- **Маркетинговые, социальные и иные исследования**
- **Информационная безопасность и защита национальных интересов страны**



Проблемы и задачи

- **Выявление стилистических, лексических, синтаксических и семантических особенностей текстов разной эмоциональной окраски**
- **Создание выборки текстов, отражающих все разнообразие настроения людей разных полов и возрастных групп, менталитета и уровня образования**
- **Исследование современных методов классификации больших объемов данных**

Проблемы и задачи

- **Исследование методов классификации текстов по тональности**
- **Проведение такого рода классификации**
- **Формирование модели, отражающей различные особенности авторского стиля сообщений в зависимости от эмоциональной составляющей**
- **Разработка системы мониторинга настроения людей в социальных сетях**
- **Ориентированность на русский язык**

Анализ тональности и методы классификации высказываний

- **Факт** — нейтральное высказывание
- **Мнение** — эмоционально окрашенное высказывание
- **Анализ тональности** — нахождение мнений и их классификация по эмоциональной окраске



Факт:
*«Это яблоко
зеленое»*



Мнение:
*«Это яблоко
восхитительно
вкусное»*

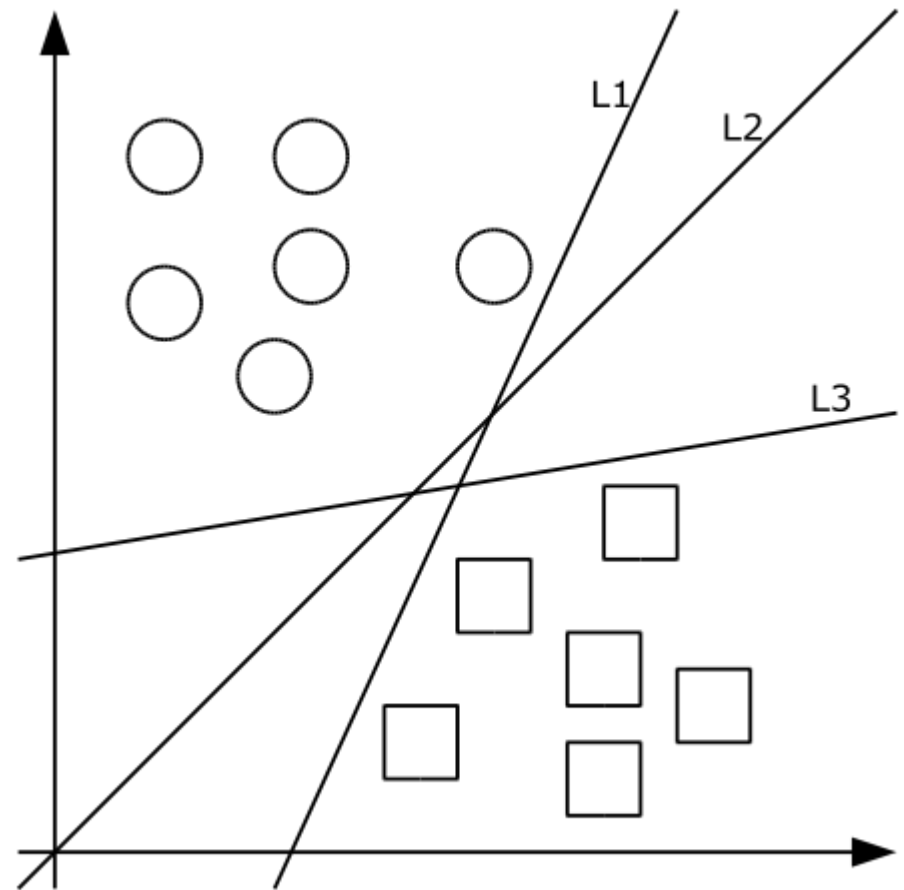
Анализ тональности и методы классификации высказываний

Алгоритмы классификации:

- **Метод опорных векторов**
- **Наивный байесовский классификатор**
- **Нейронные сети (обучение с учителем и без, походы, основанные на словарях или правилах)**

Метод опорных векторов

- **SVM, support vector machine**
- **Набор алгоритмов обучения с учителем**
- **Линейный классификатор**
- **Поиск разделяющих гиперплоскостей с максимальным зазором в пространстве**



Наивный байесовский классификатор

- Простой вероятностный классификатор
- Преимущество — малое количество данных для обучения, необходимых для оценки параметров и классификации
- Эффективно обучаем
- Теорема Байеса:

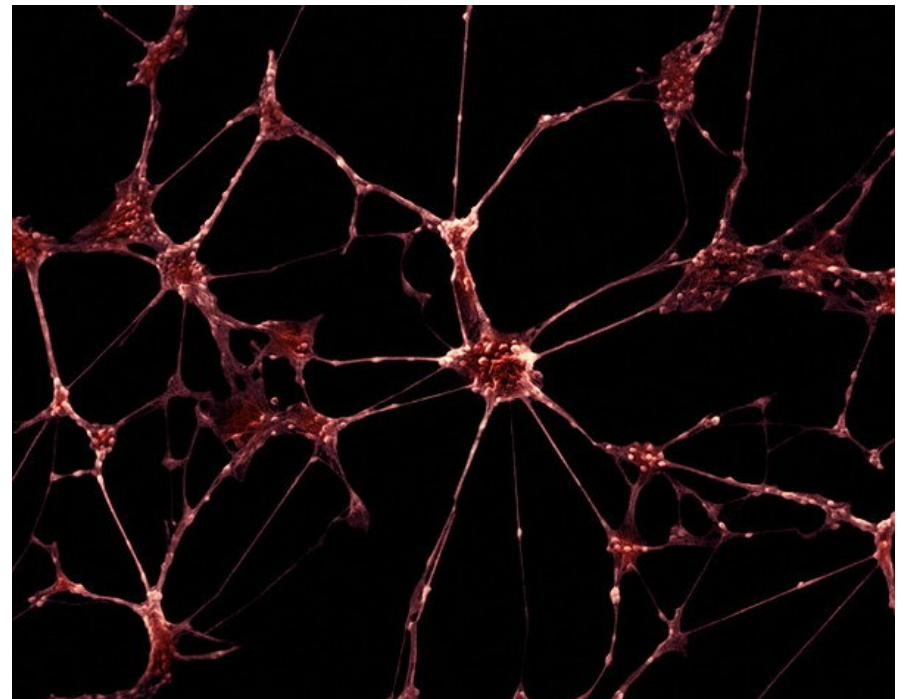
$$p(C|F_1, \dots, F_n) = \frac{p(C) p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)}.$$

Нейронные сети

Главное преимущество перед традиционными алгоритмами — возможность обучения

Подходы к обучению:

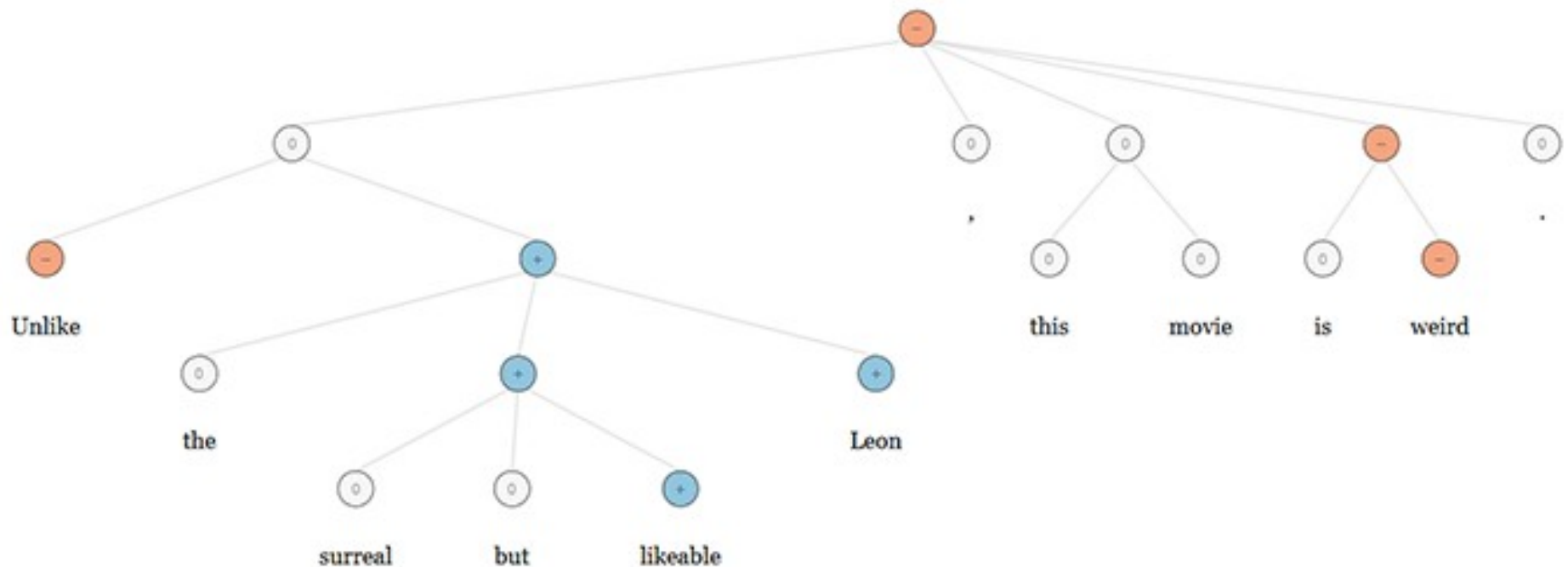
- **основанные на словарях**
- **на правилах**
- **с учителем**
- **без учителя**



Нейронные сети

Стэндфорская нейросеть

<http://nlp.stanford.edu:8080/sentiment/rntnDemo.html>

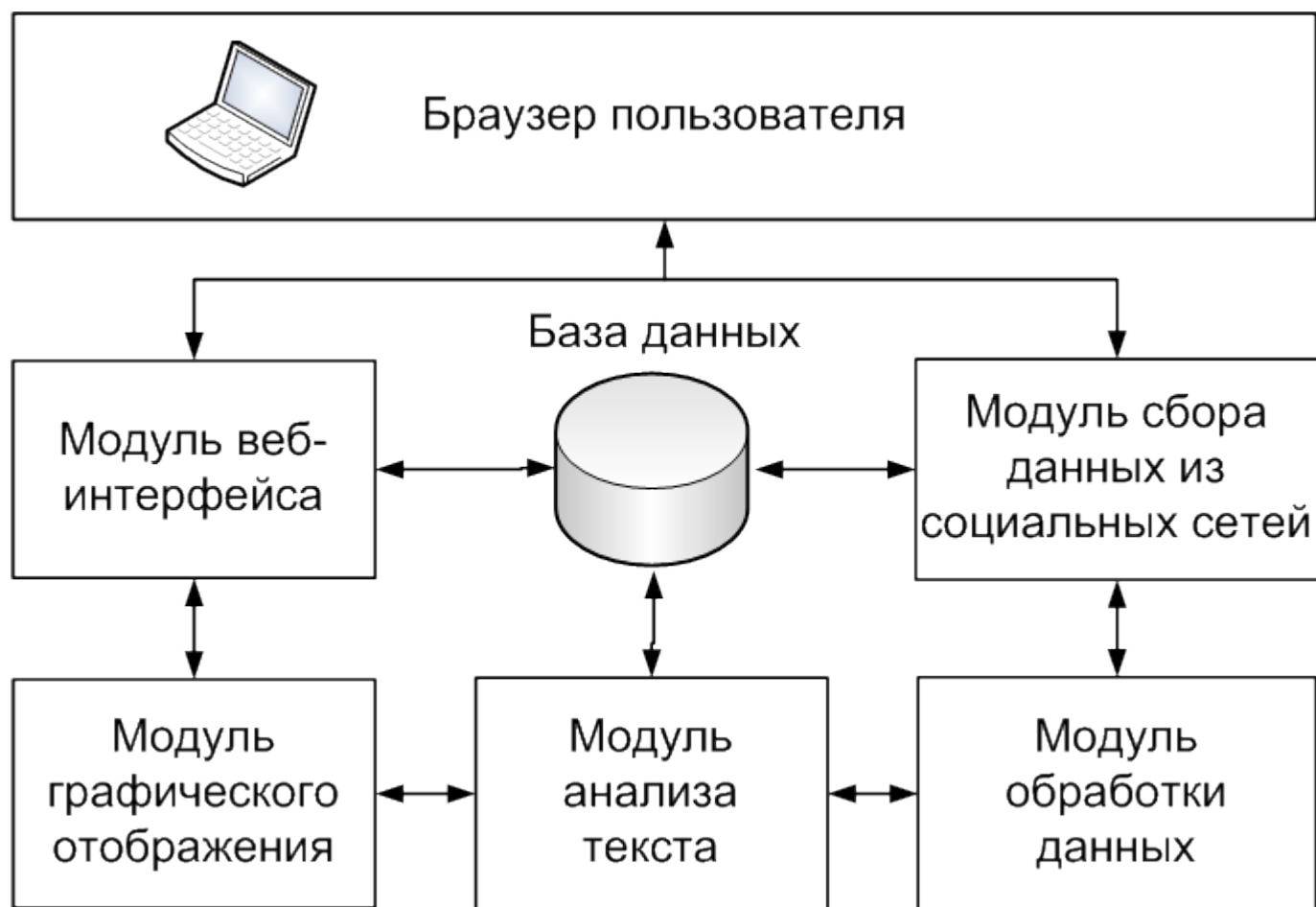


Структура разрабатываемого проекта

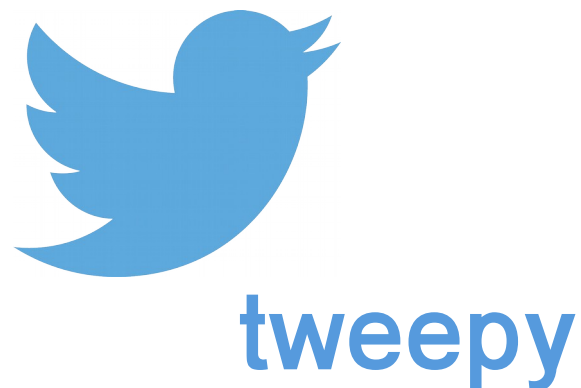
- База данных для хранения и последующего анализа данных
- Интерфейс для наглядного представления произведенного анализа
- Инструменты анализа



Структура разрабатываемого проекта

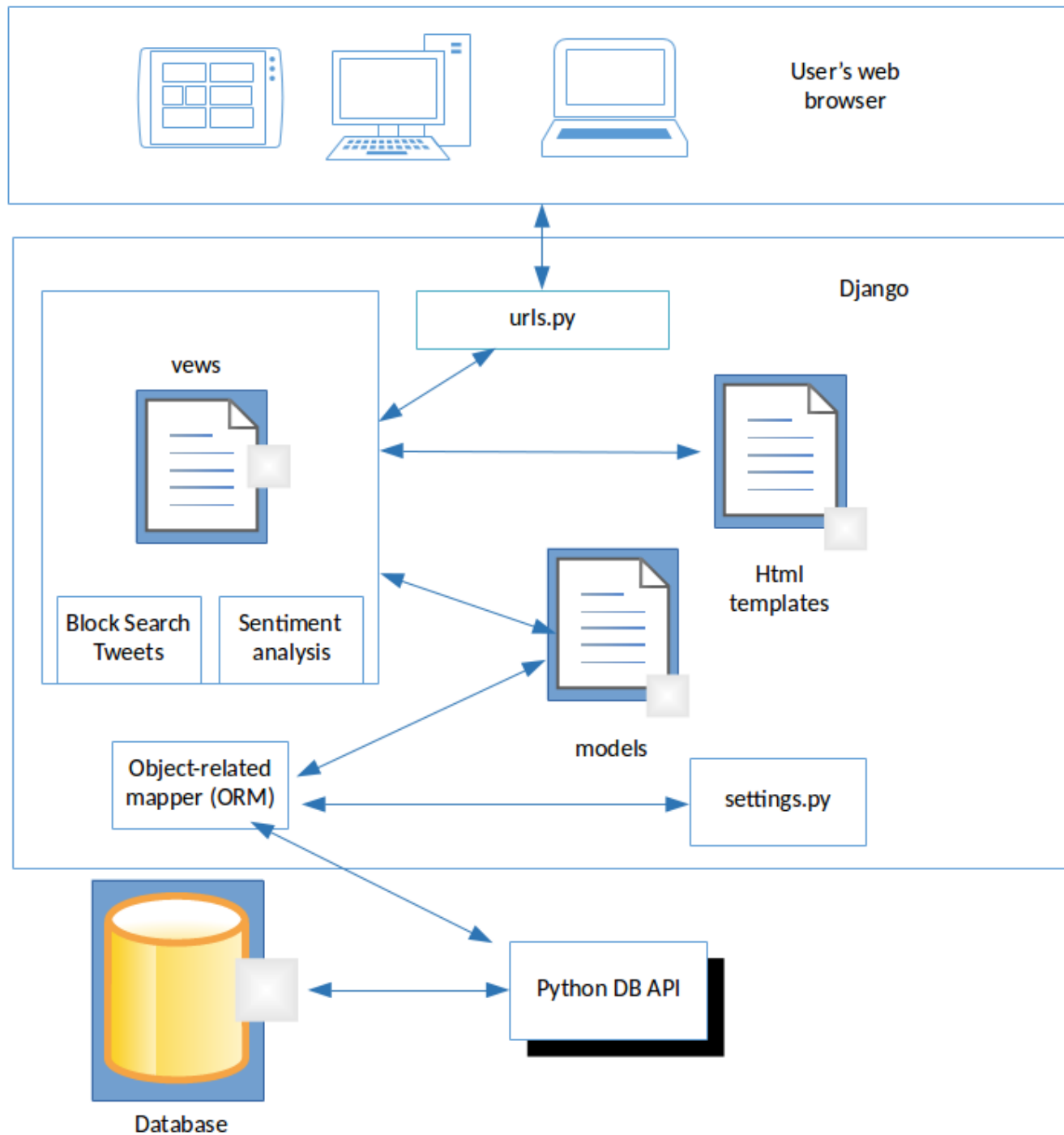


Инструменты разработки



Архитектура приложения



- **Технология ORM (Object-Relational Mapping)** — объектно-реляционное отображение
- **MVT (Model-Template-View)** — паттерн проектирования с использованием моделей, шаблонов и представлений
- **Встроенные API** для работы с БД



Результаты работы

- **Django-проект**
- **Веб-приложение для сбора твитов по определенным хэштегам и сохранением в БД**
- **Структура и модели БД**
- **Веб-интерфейс для отображения данных**
- **Исследование предметной области (разработок в области анализа тональности текста)**

Результаты работы

 192.168.1.36:8000 

Поиск

Обновить таблицу

id	text	hashtags	created at	username
639412260490858496	3 сентября. Пишу деловое письмо по КРИДО. Да, да я теперь деловая дама))) #3сентября #студентка #ТУСУР #Томск	3сентября,студентка,ТУСУР,Томск	2015-09-03 12:18:57	Irina Antipova
638777388801859584	Началась новая жизнь)отличная группа и прекрасный факультет,я рада☺ #тусур http://t.co/vTVqpb08DK	тусур	2015-09-01 18:16:12	Вероника Аверкина
638648501207965696	Отличный день. Отличная группа) #ТУСУР #Томск #1сентября2015 #студентка http://t.co/i6882LYQL8	ТУСУР,Томск,1сентября2015,студентка	2015-09-01 09:44:02	Irina Antipova
638636475278577664	Теперь я с вами #ПКФ #ТУСУР http://t.co/m4dGFHclFF	ПКФ,ТУСУР	2015-09-01 08:56:15	Maria Serebrennikova

Планы на будущее

- **Применение нейронных сетей в разрабатываемой системе**
- **Обучение системы на некоторой выборке данных**
- **Расширение возможностей системы (мониторинг других социальных сетей)**

Заключение

Наш проект можно найти здесь:

https://github.com/MarinaMeyta/social_network_analysys