

ПОСТРОЕНИЕ ЭЛЕМЕНТОВ ПРОФИЛЯ НАРУШИТЕЛЯ НА ОСНОВЕ АНАЛИЗА ЕГО ПУБЛИЧНОГО КОНТЕНТА В СОЦИАЛЬНЫХ МЕДИА

И.В. ВАЛЕНТЮКЕВИЧА

УО «Гродненский государственный университет им. Я. Купалы», факультет математики и информатики, специальность «Компьютерная безопасность», кафедра системного программирования и компьютерной безопасности

Научный руководитель - А.М. Кадан, кандидат технических наук, доцент, заведующий кафедрой системного программирования и компьютерной безопасности

Во введении рассмотрена проблема генерирования больших объемов данных и их своевременный анализ. В основной части рассмотрена реализация программного обеспечения, позволяющего строить элементы профиля на основе анализа публичного контента в социальных медиа. В заключении представлены результаты проделанной работы.

Ключевые слова: социальная сеть, социальные данные, анализ социальных сетей, машинное обучение

Введение. Важную роль в жизни людей, экономике и науке играют большие данные. В течении одного дня в нашем мире появляется больше информации, чем за всю историю человечества до 2003 года. Человечество начало генерировать огромное количество данных, что уже нет возможности оперативно обрабатывать их вручную. Их своевременный анализ является наиболее популярным направлением в современных информационных технологиях.

Основная часть. Instagram является приложением для обмена фотографиями и видеозаписями с элементами социальной сети, позволяющее снимать фотографии и видео, применять к ним фильтры, а также распространять их через свой сервис и ряд других социальных сетей.

В ходе исследования социальной сети были выделены следующие основные сущности: **User** (пользователь), **Media** (публикация), **Caption** (текст публикации), **Category** (категория), **Location** (локация), **Like** (лайк), **Label** (объект, изображенный на фото). Связи и атрибуты между сущностями приведены на рисунке 1.

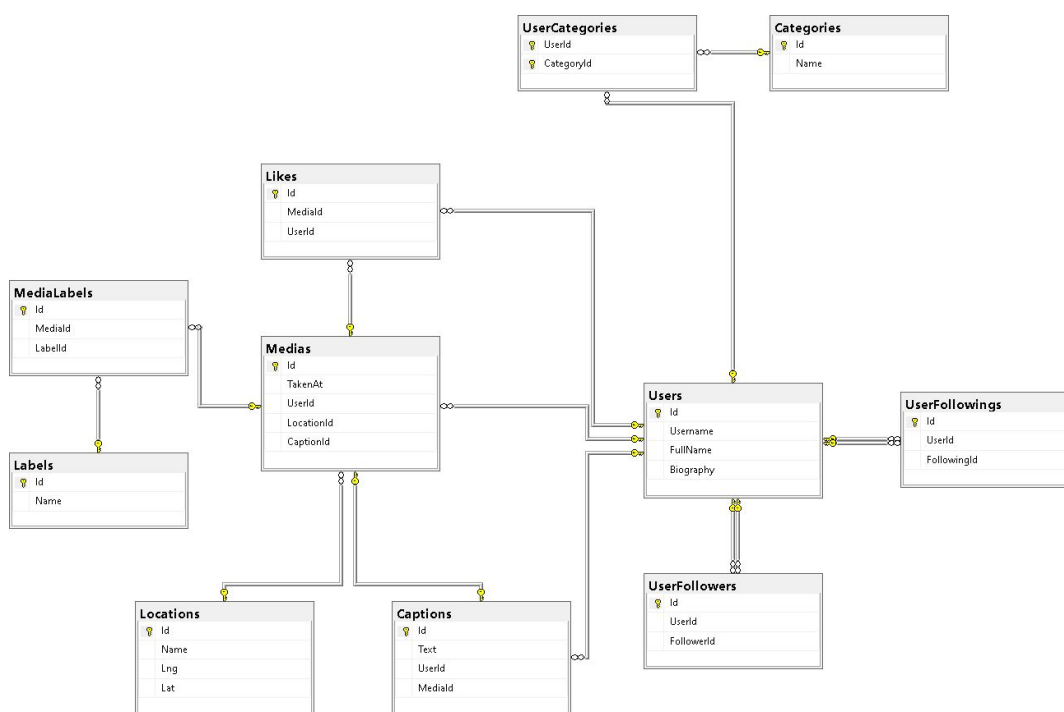


Рисунок 1 - Диаграмма таблиц в БД

На основании данных моделей было разработано клиент-серверное веб-приложение, которое позволяет: загружать данные о пользователе из социальной сети Instagram, включая все его публикации, подписки и подписчиков; идентифицировать объекты на фото; производить анализ контента, аудитории и лайков пользователя; добавлять категории аккаунтам и осуществлять поиск связей между ними.

Для разработки серверной части приложения была выбрана платформа .NET Core 2.2. Основными преимуществами данной технологии является кроссплатформенность (Windows, Linux, macOS), встроенное введение зависимостей, упрощенный процесс веб-разработки.

Для удобного обращения к API социальной сети Instagram используется библиотека InstaSharper Api.

Для идентификации объектов на фото использовалась облачное решение Vision AI от Google. Входными данными является изображение одного из популярных форматов (JPEG, PNG). На выходе сервис отправляет JSON с набором меток и их совпадениями. Так, на рисунок 2 сервер прислал ответ представленный в таблице 1.



Рисунок 2 - Пример изображения

Таблица 1 - Пример ответа Google Cloud Vision AI

Метка	Вероятность совпадения, %
People In Nature	98
Sitting	96
Photograph	95
Green	94
Rock	93

В приложении реализована возможность создать и назначить категорию пользователю. Для того, чтобы найти аккаунты, которые относятся к данной категории, необходимо создать базис профилей, точно принадлежащих определенной категории. Например, выделим и создадим новую пользовательскую категорию “Студент КБ-161”. После этого нужно найти как минимум 3 аккаунта, которые точно принадлежат данной категории. Далее, программа автоматически проанализирует подписчиков и подписки аккаунтов, найдет их пересечение, и сформирует новый базис аккаунтов, которые с определенной вероятностью относятся к студентам КБ-161. Это возможно благодаря тому, что люди подписываются на знакомые профили, которые соответствуют их интересам.

Для разработки клиентской части использована библиотека ReactJS. Данная библиотека позволяет создавать сложные интерактивные интерфейсы основанные на компонентах. Пример созданного интерфейса представлен на рисунке 3.

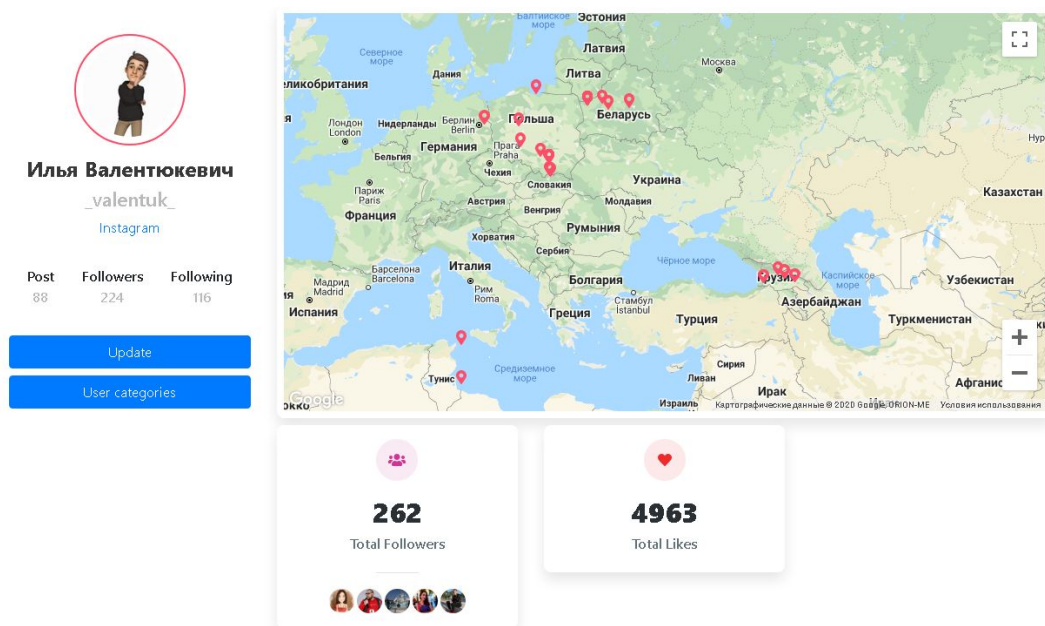


Рисунок 3 - Интерфейс приложения

Закключение. Таким образом в ходе работы было разработано программное обеспечение, позволяющие формировать элементы профиля на основе анализа публичного контента в социальных медиа.