

Сбор гео-данных

Мотивация

- Среднее время в пути для учеников Москвы?
- Автоматизация доставки и сбора товара
- Визуализация карт походов, забегов
- Соотношение между следами гео-следов приборов

Основные понятия

- *Geocoding* – нахождение географических координат по названию объекта;
- PointOfInterest(POI)
- Polygon
- GeoJSON
- GeoDataFrame

Open Street Maps

- Открытое движение
- Многие университеты в России постарались, чтобы карты были достаточно точны и регулярно обновлялись
- OpenStreetMaps имеет сложную систему хранения данных, которую каждый воспринимает по своему. Их можно использовать только для оценки

Модель данных

- Модель хранения графовая – узлы, ребра и их объединения
- Каждый из элементов может иметь тэги. Они как правило имеют типовые названия и служат для поиска:
 - amenity=restaurant
 - name=Example Restaurant
 - cuisine=Italian
- Правило назначения тэгов описано в документации OSM, так тэг [admin level](#) может иметь лишь 9 значений в зависимости от уровня правительственной иерархии

Nominatim

- API позволяющий бесплатно запрашивать данные;
- Сервис очень мощный, вы можете выгрузить несколько гигабайт данных за раз;
- Не подходит для коммерциализации из-за отсутствия быстрых запросов, но можно поднять свой [сервер](#);
- В совокупности с Python библиотекой [osmnx](#) позволяет представлять данные в виде графа для анализа

НЗ

- Гексоганальная система с иерархическими уровнями
- Прекрасно подходит для подсчета статистик по разреженным данным

Folium

- Открытые веб-карты в виде JavaScript компонент с возможностью кодовой модификации.
- Информация для визуализации карт загружается клиентом из Nominatim, но можно добавить и другие источники
- Модель данных выполнена в виде системы слоев, которые позволяют наносить точки и регионы на карту
- Модель общения GeoJSON

GeoIP

- Провайдеры отправляют сетевые маски и местоположение провайдера в публичные сервисы;
- Если пользователь поделился (=подключился к вам) IP, вы можете узнать его примерное местоположение;
- VPN-ом на постоянной основе сейчас пользуются примерно 40% жителей России (~80% платежеспособной аудитории), что осложняет коммерческое применение технологии;

GPS/Cellular triangulation

- Мобильные телефоны используют [Assisted GNSS](#) технологию для поиска своего положения. Это технология очень точна на открытых территориях (~0.5 м). GPS также работает с точностью до полуметра.
- Измерение дистанций в закрытых помещениях сложно из-за отражений сигналов. Обычно используются опорные [Wi-Fi](#) точки

Платные сервисы

- DaData – лучший геокодер в РФ;
- 2Gis – лучшие карты + сервис с статистикой по продажам;

Выводы

- Открытые решения имеет смысл применять для разведки и визуализации данных;
- Если используете карты в коммерческих задачах, то купите подписку. Пользователю приятнее доплатить 10 рублей вместо поиска курьера/брождения по району;

Практика

Домашнее задание

Полезные ресурсы

- На русском языке с кусочками кода
<https://habr.com/ru/articles/579838/>