Challenges

Traveling is a very desired activity these days, but it is also very expensive.
Humanity found creative ways to make traveling affordable, so many people will be able to enjoy vacations often without tearing their pockets.
One of these awesome ideas is sublet. It makes so much sense - save your vacation budget by staying at someone else's apartment, instead of a pricey hotel - a win-win situation for everyone.

Motivation

In Israel, it might be exhausting to look for the right sublet :
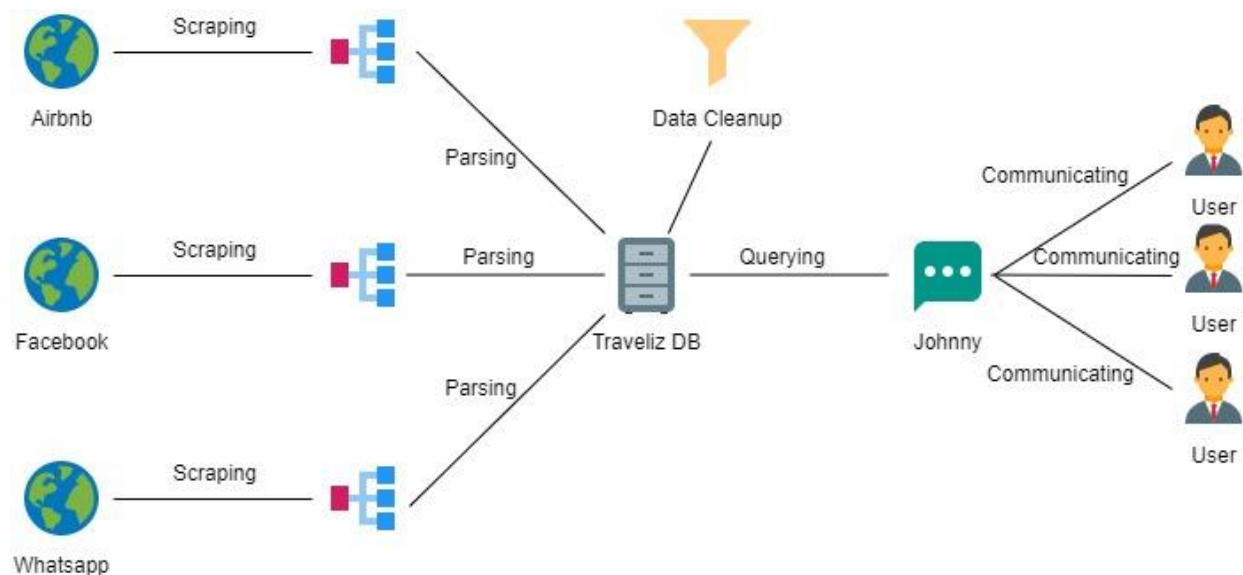- Sublets are splitted between different platforms such as : Airbnb, Facebook groups and Whatsapp groups.
- Sublets posts mainly based on free-text, which makes it impossible to filter on, and search wisely and according to needs.

Solution

Traveliz - an online Cross-Platform Database for sublets. We use technology to collect data from Airbnb, and multiple Facebook/Whatsapp groups.

The way to communicate with our system is with Johnny, a telegram-bot. He knows to discuss with you about the sublet details - just give him dates, location, budget and number of rooms and he will take care of your request.

Architecture

Challenges

1. **Scraping in large scale**
   Scraping data from the mentioned platforms requires handling with different kinds of protections and limitations.
   To overcome this, we used various types of techniques such as : IP rotation, using specific user agents, user switches, etc.

2. **Parsing free-text to metadata**
   Extracting metadata (dates, price, location, rooms, etc.) from free-text posts is challenging. Our system is basically scanning each post, and tag words under labels, following the combination of three different methods :
   - Regex parser : identify label's patterns and tag accordingly (suitable mostly with dates, but relevant for rooms and price as well).
   - Location parser : crosscheck words with a predefined DB of locations in Israel, and tag accordingly.
   - NER (Name Entity Recognition) : using Machine learning model to tag words.

Location	Start Date	End Date	Price		Rooms	
Tel Aviv	14.4.22	24.4.22	1,500	Per period	2	Whole Apt.
Tel Aviv	18.5.22	18.7.22	2,700	Per month	4	Shared Apt.

סאבלט פסח
דירת 2 חדרים משופצת עם גינה קטנה, שקטה ובמיקום נדיר במרכז פלורנטין!
מרחק דקה הליכה מהפאבים, מסעדות ובתי קפה.
בין התאריכים 24.4-14,
1,500 ש״ח כולל.
פרטים נוספים בפרטי.

סאבלט בצפון הישן (רחוב פרופסור שור) ✨
2 שותפות, מחפשות שותפה שלישית
כניסה ב-18/05 למשך חודשיים/שלושה
2700₪

3. **Data cleanup**
   Cleanup irrelevant records due to :
   - Sublets which were already taken. Could be done by parsing comments or checking the existence of a pre-parsed post.
   - Posts published by people who are looking for sublets.

Code

https://github.com/avishayAv/traveliz/