Project Name: **A Web scraper**

A Web Scraper That Scrapes Company Information By Leveraging Google + Facebook.

THE PROBLEM

Client needed something really cool to scrape company information AKA collect leads by niche and the location specified by software. So that they are looking for a better solution but unfortunately most of their output result were very bad! So they badly wanted a solution so that they needed to take step forward for this solution.

PROPOSED SOLUTION

So we got their idea and proposed a solution for a chrome extension plugin software. To use with Chrome browser. Because they wanted something that maybe saves financially as well. So that's we proposed to client for costs, timeline, and the use-case in short as well.

PROJECT STATEMENT

In Project statement both party the client and the developer team clarifies and notifies about the deliverables, costs, and timeline for this project. And also more few things has been taken care of.

RELATED PROJECTS

- Data Scraper - Easy Web Scraping (https://chrome.google.com/webstore/detail/data-scraper-easy-web-scr/nndknepjnldbdbepjfgmncbggmopgden)

- Web Scraper - GetData.IO (GetData.IO)

REQUIREMENT GATHERING

Client discussed with us about the problem and then we told them that this can be achieved. And afterwards when the deal is final we will have to take the requirements officially.

Below few points we have noted.

- ▶ Functional requirements
- ▶ Timing
- ▶ Marketing, Sales & Service requirements
- ▶ Technical Requirements
- ▶ Look & Feel requirements

- Performance Requirements
- Operational Requirements
- Etc

DOCUMENT READING

- 1. Business Case
- 2. Project Charter
- 3. Project Management Plan
- 4. Project Schedule
- 5. Project Status Reports
- 6. Project Budget Tracker
- 7. Project Closure Document (What the project delivered)

OBSERVATION

Where applied?

▶ Collect Requirements

▶ Monitor Communications

▶ Manage Stakeholder Engagement (is used to stay in touch with the work and attitudes of the project team members and other stakeholders)

QUESTIONNAIRE

- ▶ What Technology Do you Prefer?
- ▶ What type of authentication system would you like to have?
- ▶ What is your preferred time to get that project?
- ▶ Do you agree that project planning techniques are important to avoid cost or time overruns?

INTERVIEW

After gathering all the requirements and like that paper a interview is required to discuss about everything and make sure client and the project manager both are on same page. So that unwanted gap and misconception can be overcome by having like that meeting.

This will help both parties to have a clear view.
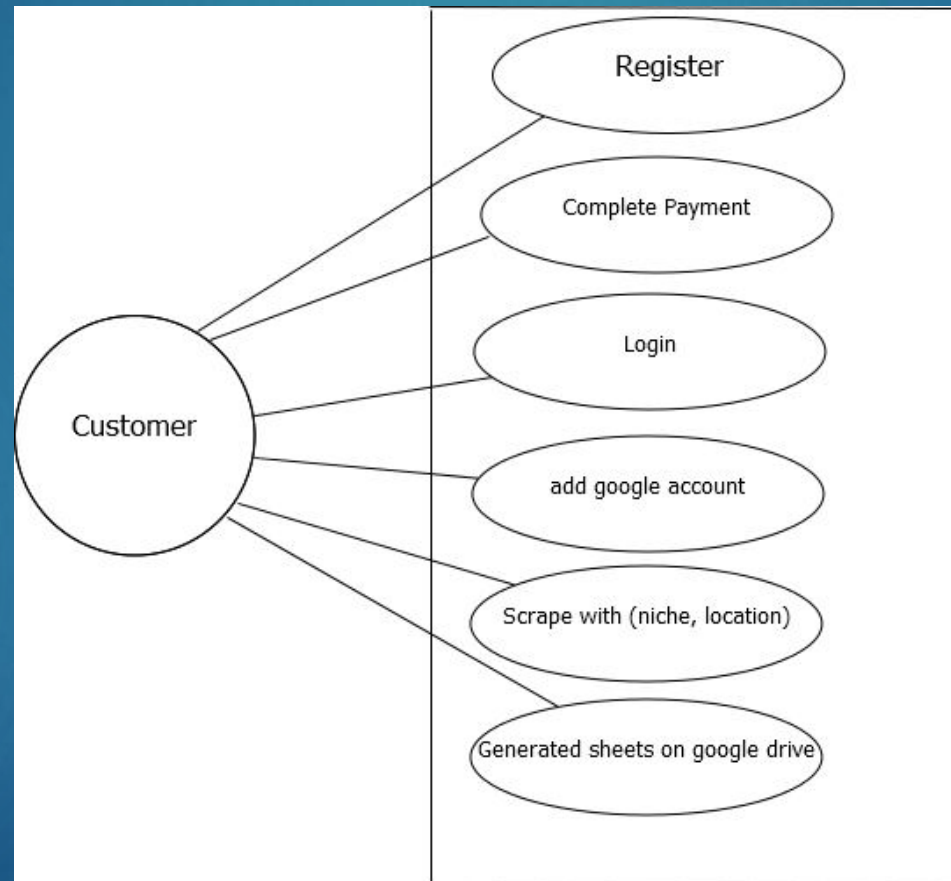
FUNCTIONAL REQUIREMENTS

We took all the functional requirements here for the project:

- User registration

- User should have a function to complete his subscription

- Users need to be authenticated with google to get the permission for saving scraped results on google drive/sheet.

- Users should be able to search by niche and location as well. Where users can upload a list of location to search on each of them.

- When the hit submit to scrape it should automatically complete all the necessary steps and then generate a spreadsheet on google drive/sheet for the permitted user.
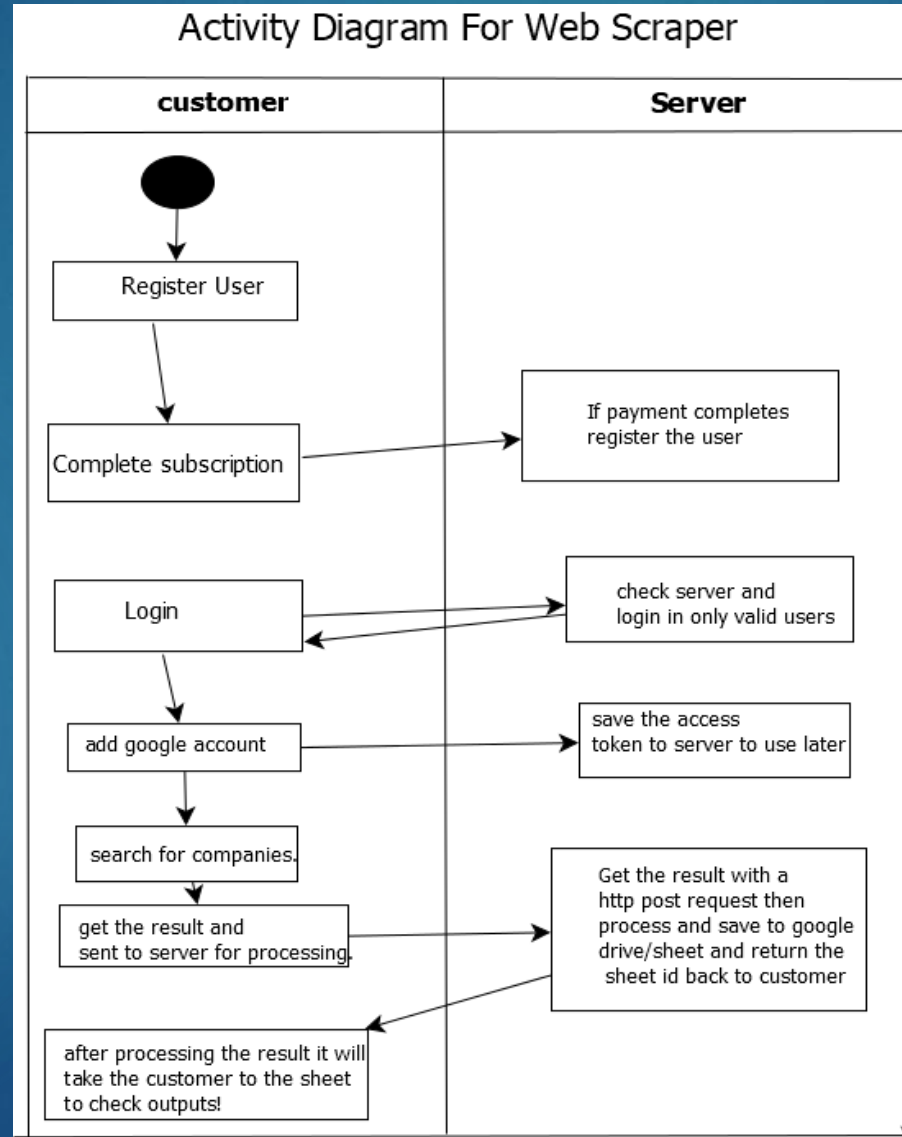
NON FUNCTIONA REQUIREMENTS

- If the user's payment date expired they should be blocked to using this software. Then they should get a way to renew their license

- If the user didn't add their google account with this software then they shouldn't be able to start the scraping. And ask them to do so

- If scraping taking longer time then expected then automatically it should close and clear stuffs.

- System must need to have proper effectiveness, and that's need to be make sure. Also the maintainability should be easier.
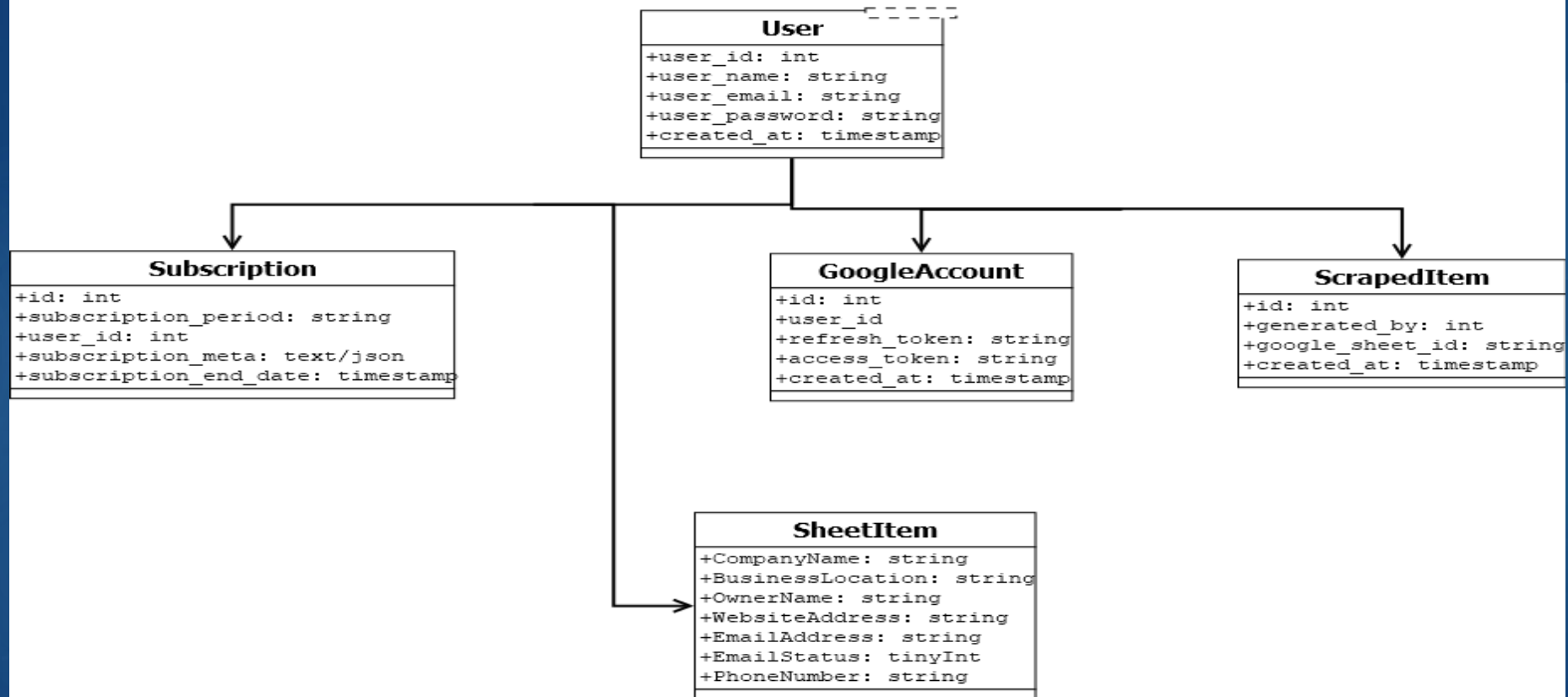
USE CASE DIAGRAM

ACTIVITY DIAGRAM

Activity Diagram For Web Scraper

CLASS DIAGRAM

Class diagram for web scraper

User
+user_id: int
+user_name: string
+user_email: string
+user_password: string
+created_at: timestamp

Subscription
+id: int
+subscription_period: string
+user_id: int
+subscription_meta: text/json
+subscription_end_date: timestamp

GoogleAccount
+id: int
+user_id
+refresh_token: string
+access_token: string
+created_at: timestamp

ScrapedItem
+id: int
+generated_by: int
+google_sheet_id: string
+created_at: timestamp

SheetItem
+CompanyName: string
+BusinessLocation: string
+OwnerName: string
+WebsiteAddress: string
+EmailAddress: string
+EmailStatus: tinyInt
+PhoneNumber: string

CONCLUSION

So this was it! We've tried to merge everything about this project in this presentation.