Assignment

1. A developer is assigned a task to scrape 1 lakh website pages from a directory site, while scrapping he is facing such hcaptcha, which are placed to stop people from scrapping As a project Coordinator suggest ways to solve this problem

ANS:

As a project coordinator, I would suggest the following ways which I used to solve the problem of hCaptcha while scraping 1 lakh website pages from a directory site:

- **Rotating IP addresses**. One can use different IP addresses for each request that can make to the directory site.

- **Using a CAPTCHA resolver service/bot**. Using this will save my time.

- **Using proxy server**. A proxy server can hide real IP address from the directory and chances of being blocked.

- **Slow down your scraping.** Scraping too quickly can trigger the hCaptcha. It is important to slow down your scraping to make it look more like human behavior.

---

2. Our client has around 10k linkedin people profiles, he wants to know the estimated income range of these profiles. Suggest ways on how to do this?

ANS:
There are a number of ways to estimate the income range of LinkedIn profiles. Some of them includes :

- **Using a LinkedIn salary estimator tool.** This is the majorly used in these types of tasks There are a number of others tools available that can estimate a person's salary based on their job title, experience, company, and location.

- **Looking at the person's job title and experience**. Certain job titles and levels of experience are associated with their income.

- **Looking at the person's company size and industry**. Larger companies pay their employees more.

- **Looking at the person's location.** The cost of living in different places can vary significantly, so this can also affect income.

---

3. We have a list of 1L company names, need to find linkedin company links of these profiles, how to go about this?

Submitted by : AMAN

ANS:

To find LinkedIn company links for a list of 1L company namesI would use use a combination of the following methods:

- **Will use the LinkedIn search feature**. I will search for LinkedIn company pages by name.

- **Make the Use the LinkedIn API**. I can use the API to search for LinkedIn company pages and to get their URLs.

- **Use a third-party tool.** Third party tools can also be taken into consideration.

---

 4. How to identify list of companies whose tech stack is built on Python. Give names of 5 companies if possible, by your suggested approach

ANS:

To identify a list of companies whose tech stack is built on Python, one can use a combination of the following methods:

- **Looking at the company's website.** Many companies list their tech stack on their website.

- **Looking at the company's job postings.** Job postings often list the technologies that the company uses.

Use a third-party tool. There are a number of third-party tools that can help you to identify companies whose tech stack is built on Python.

There are many companies whose tech stack is build on python such as **FACEBOOK, NETFLIX , GOOGLE , DISCORD, UBER** and many more

---

5. Need to find an API, through which we can send linkedin messages to other linkedin users

ANS:
There are a number of APIs that can be used to send LinkedIn messages to other LinkedIn users. Few of them are:

- **LinkedIn Sales Navigator:** It is paid tool that can be used to text people who are not in your connection.

- **Phantombuster**. Its is also a paid tool that is used for automation for sending messages.

- **Expandi.** Like Linkedin sales navigator it is also used to send messages to people who are not in your connection.

Submitted by : AMAN