

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 suggests ways to solve this problem.

⇒ There are different ways and methodologies we can work with to deal with captcha while scraping data from websites:

- Slowing down the scraping to make my program behavior less robot / automated-like.
- Make use of proxy servers to minimize IP address tracing
- We can solve the captcha manually every time we face a captcha, though in our case I don't think that will be a good choice.
- We can use pre-build web-scraping tools like " selenium " and "puppeteer".

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?

⇒ We can identify the companies the profiles are working on from linkedin and then we can get insights from popular salary estimator tools(eg: Glassdor, PayScale, Indeed) about the company and the job positions salary. Or we can build machine learning models for the same purpose.

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

⇒ There are multiple Linkedin company directories available we can use them or Linkedin itself provides services.

We can use Linkedin automation tool like- "phantombuster", in this process we need to specify the list of companies which we already have then it will automatically find links for each company.

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.

⇒ we identify this in multiple ways :- 1. Scape the tech stack data from the company websites.

2. Databases available online to find out tech stacks of the companies.

3. Using job search engines

5 companies using Python in their stack - Dropbox, Spotify Google, Reddit,Uber

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

⇒ There is no official API available as of now to send message , but we can use third party API like - “Phantombuster” or other API which are available in the market, though we need to be careful while using these API because linkedin has Strict terms and conditions regarding API's.