

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.

Solution:

The reason a website has captcha is to prevent machines or frequently appearing IPs from Scrapping, So we can avoid this captcha issue by rotating IP address to avoid being detected or blocked.

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?

Solution:

For finding estimated income range of these profile, we can leverage the company name from user linkedin profile, then we can visit company website and see the potential income ranges based on job titles.

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

Solution:

We can use Web Scrapping tools like SCRAPPER, Python libraries like BeautifulSoup to extract links by iterating through company names and fetching the links or can be done manually.

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.

Solution:

We can check job opportunity section on company website where the required skill can be their tech stack build on or by searching the company on linkedin and browse their profile and many times it is mention in their about section and also app like Glassdoor which includes discussions and reviews about company that might reveal the technologies used by company.

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

Solution:

To find Linkedin API, We will need to obtain the appropriate API permission, so first authenticate API request for that it requires linkedin developer account, once we have obtained the access we can post request to the /v2/messages with the recipients linkedin id and message text in the request body.