

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 - Ways that we can use to solve the problemes -

Emulate Human Behavior:

- **User-Agent Header:** Set your User-Agent header to mimic a popular web browser
- **Rate Limiting:** Introduce delays (using `time.sleep()`) between requests to avoid appearing as a bot.
- **Randomize Interactions:** Simulate interactions with the website in a less systematic manner, such as clicking on links and scrolling.
- **Session Persistence:** Use libraries like ``requests`` to manage sessions and handle cookies.

Cookie Handling:

- **Store and Send Cookies:** Send an initial request to gather cookies, store them, and include them in subsequent requests to maintain session context.
- **Cookie Management:** Respect the website's privacy policies and terms of service when handling cookies.

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- We can use the LinkedIn salary insights and salary estimation tools for this one

To use the salary Insights feature of LinkedIn we must have linkedin premium account.

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

Ans. - There are a number of LinkedIn company finder tools available online. we can go for that.

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- We can use the tech stack search engine available online to filter the companies based on the tech stack.

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

Ans- There is no official LinkedIn API for sending messages to the other LinkedIn users but we can write our own Python script for doing that