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

Solution : - As Project Coordinator there are multiple ways to handle this situation .

* Firstly we can increase the wait time between scraping request
* do not send mass request to the server.
* Rotating IP addresses because Captchas usually appear when you're using blacklisted IPs. You can try rotating your IP address, especially if you're using a quality proxy network.
* Save cookies because Cookies are small pieces of data that are stored on the user's computer by a website. H captchas often use cookies to track the user's activity and identify them as a human.
* Avoid hidden traps ; H captchas often include hidden traps that are designed to catch bots
* Hide any automation indicators h captchas can be used to detect automated traffic by looking for certain indicators from the same Ip address.
* While scraping such huge number of pages from the site I personally suggest that we can change scraping methods frequently so it does not get any issues

And project won’t be at hold

* Lastly we can use h captcha resolver ; and there are a number of captchas resolvers available that can solve h captchas automatically. This can be a good option for developers cause while scraping such huge number of pages they do not want to spend time manually solving Captchas.

1. **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 : We can use the Linkedin’s own Tool of LinkedIn Salary Insights :

* LinkedIn Salary Insights is a tool that provides users with salary estimates for different job titles, companies, and locations. To use this tool, you would need to create a LinkedIn Talent Insights account and generate a talent pool report for your 10k LinkedIn people profiles. The report will include a compensation summary, which will show the average total compensation and the average compensation range for your talent pool.

Use a third-party salary estimator :

* There are a number of third-party salary estimators available, such as Salary.com and PayScale. Or Glassdoor.com These websites allow you to enter a job title, company, and location to generate a salary estimate. You can use these websites to estimate the income range of your 10k LinkedIn people profiles one by one.

Machine Learning and Data Analysis :

* Consider employing machine learning models or data analysis to mine patterns and correlations between available profile data and known income levels in similar job roles or industries. This approach might be more complex and require data science expertise.
* As per project coordinator I can suggest not to use the complex approach towards the solution unless and until it requires accuracy.

Use a combination of methods :

* Look for profiles that list their current and previous salaries. This is the most accurate way to estimate their income range.
* Look for profiles that list their job title, company, and location. You can use this information to research the salary range for similar positions.
* Look for profiles that list their skills and experience. You can use this information to infer their income range.
* Look for profiles that have been endorsed by others for their skills and experience. This suggests that they are highly valued in their field, which may indicate a higher income range.

Experience and Seniority:

* Profiles often mention the years of experience. Seniority and experience can be indicators of higher income. However, it's essential to note that experience doesn’t always correlate directly with income, especially in diverse fields and job roles.
* **With the above Solutions I can asure we can definetly give estimated income range of these profiles for Our client requirement.**

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

Solution : Manual Search :

* + - Manually search for each company on LinkedIn using their search bar. This might be time-consuming but can be effective for smaller lists. Simply type the company name into the search bar on LinkedIn and look for the company page.
    - For this number of company names I do not suggest the manual Search.

Company Page URL Format:

* + - LinkedIn typically follows a standard URL format for company pages. It's usually structured as www.linkedin.com/company/[company-name].
    - So With URL format we can structure the company links list with help of

Excel Formating.

With the help of some tools :

* + - Botster: Enter company names or domains on the LinkedIn Company Finder's start page. The URLs will be exported into a spreadsheet.
    - PhantomBuster: Create a free account, enter the company names, and specify the number of names to process per launch. You can download a CSV spreadsheet or JSON file of all your LinkedIn company pages.
    - TexAu: Create a free account and make a list of companies in a Google Sheet.
    - Apify: Feed it a list of companies you are interested in.

Use a bot :

* + - to find LinkedIn company page URL links in bulk, then go to the LinkedIn Company Finder's start page, enter company names (or domains), and launch the bot. The necessary URLs will be exported into a spreadsheet.

API :

* + - LinkedIn API to search for and retrieve company pages. You'd need to register your application on LinkedIn's Developer portal to obtain access and authorization to use their API.
    - Some of the API requires permission and Subscription
    - The Free API Python installation link - pip3 install git+https://github.com/tomquirk/linkedin-api.git

Web Scrapping :

* + - web scraping using programming languages like Python and libraries such as BeautifulSoup or Selenium to automate the search process. This involves writing a script that searches LinkedIn for each company name and extracts the company page URLs.

1. **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 :

* + - Identifying the tech stack of companies, especially specific programming languages like Python, can be challenging as this information is not always publicly disclosed.
    - According to the Stackshare page for Python, there are 6,200 companies using Python in the real world.

Job Postings Analysis :

* + - You can use a search engine like Google to search for companies that mention Python in their website content or job postings. This will give you a list of companies that are likely to be using Python in their tech stack. Many companies advertise job openings that require Python skills. Websites like LinkedIn, Glassdoor, Indeed, and company career pages can be searched for job postings that mention Python. This can indicate companies using Python in their tech stack.

Technical Conferences and Events :

* + - Companies sometimes present or sponsor events related to their tech stacks. Checking speakers or sponsors at Python-related conferences or events could hint at companies using Python.

Search Engine Or AI tools :

* + - Google
    - Facebook
    - Amazon
    - Netflix
    - Spotify

Use a tech stack identification tool : Some popular tech stack identification tools include:

1. BuiltWith
2. Wappalyzer
3. WhatRuns

company's open source contributions :

* + - If a company is using Python in their tech stack, they are likely to be contributing to Python open source projects. You can search for the company's name on GitHub and other open source repositories to see if they are contributing to Python projects.

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

Solution :

* + - As per Knowledge Linkedin Does not allow non premium user to send messages to other users without connection request LinkedIn doesn't offer a public API endpoint that allows sending direct messages to other LinkedIn users outside of the official LinkedIn platform or applications. LinkedIn's API has limitations in regard to messaging functionalities due to privacy and user protection policies.