**Introduction:**

We propose to develop an automated system capable of generating LinkedIn user profiles, establishing connections, and analyzing data for targeted purposes. This system will leverage various tools and APIs to streamline the process and ensure efficiency.

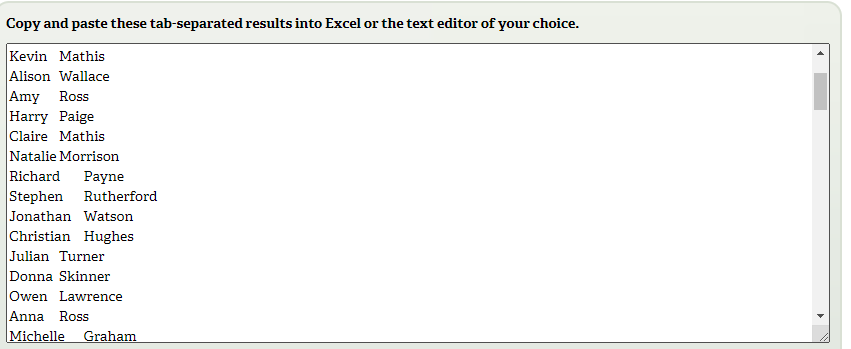
**Project Steps:**

1. **Proxy Management:**

* Procure residential proxies from Bright Data’s database (estimated cost ~ $70/mo for 84 proxies and $8.40/GB of traffic assuming ~ 3 GB of traffic per month).
* Implement a functioning proxy rotation system to efficiently utilize proxies and avoid detection.
* Develop a mechanism to store cookies for each proxy/webdriver instance to maintain session persistence.
* Integrate the 2Captcha API to handle captchas effectively. (<https://2captcha.com/pricing>)

1. **Name Generation:**

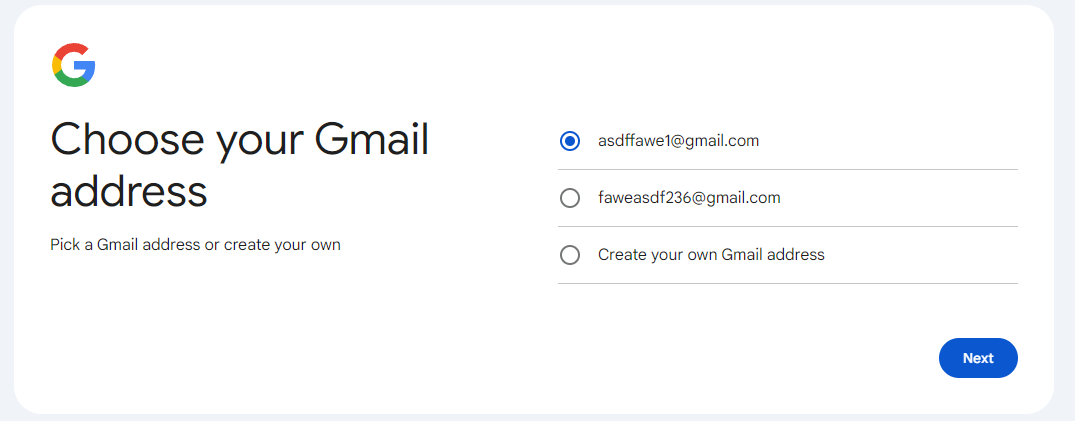
* Generate and scrape 100 fake names from <https://homepage.net/name_generator/> for LinkedIn profiles. (Example below)



* Save the generated names to a Postgres database for future reference.

1. **Gmail Account Creation:**

* Automate the creation of Gmail accounts using the generated names.
* Extract and store the first Gmail address suggestion.



* Scrape passwords from 1Password and securely save login information to a database. (<https://1password.com/password-generator/>)



1. **Phone Number Verification:**

* Utilize a free TextFree API (<https://github.com/JJTech0130/textfree>) to generate disposable phone numbers for Gmail verification.
* Register the Gmail accounts using the phone numbers and retrieve access codes using string manipulation.

1. **LinkedIn Account Creation:**

* Create 100 LinkedIn accounts by signing up with Google accounts.
* Populate fake profile data including education and work history (excluding government agencies).
* Fetch fake profile pictures from This Person Does Not Exist (<https://thispersondoesnotexist.com/>).
* Establish at least 150 connections with individuals at non-government companies, with a preference for current organizations.
* Gradually increase connections over several weeks to build account trust.

1. **Data Collection:**

* Automate doomscrolling on the CMS LinkedIn page to capture names of people and store them in a database (~5000 names).

1. **Scaling and Optimization:**

* Activate proxy rotation and run 12 simultaneous instances, each with a different LinkedIn account.
* Rotate proxies and LinkedIn accounts periodically to avoid detection and populate the database with user profile data.

1. **Data Analysis:**

* Clean and analyze the collected data.
* Utilize natural language processing techniques to sort profiles by skills and other relevant criteria.

1. **Web Application Development:**

* Create and host a website using React/Django, hosted on AWS EC2.
* Allow users to search by specified input fields.
* Provide options to export search results as CSV, XLSX, or view them in an interactive table format.

**Conclusion:**

The proposed project aims to develop a comprehensive system for automated LinkedIn user generation, data analysis, and user-friendly access to search functionalities. Through the integration of cutting-edge technologies and APIs, we are committed to delivering a robust solution that effectively meets the project requirements and provides actionable insights for strategic partnerships. This system will facilitate targeted analysis to identify potential partners or collaborators based on the contracts CMS employees are involved in, thereby enhancing the project's utility and value. Our goal is to create a powerful tool that streamlines LinkedIn user generation, data analysis, and facilitates strategic decision-making processes.