

CS Project Overview – LinkedIn Scraper Tool

Andrew Jerome, Soham Desai, Arik Gitman, Darsh Patel, Dylan Shah

Professor Kahanda

April 11, 2025

---

# Rutgers Alumni LinkedIn Scraper

*A Simple Overview of the Program and Its Features*

---

## What is it?

This project is designed to help identify accurate LinkedIn profiles for Rutgers alumni. By entering a person's name and graduation year, the program searches the internet and returns the best match. It's especially useful for analyzing career paths, work locations, and estimating income data.

---

## Manual Search Support

Users can manually input a name and graduation year to narrow down results. This is helpful when only one person is being searched or when more accuracy is needed.

---

## Works Even Without Google Search Tools

If the primary tool (SerpAPI) is unavailable due to rate limits or errors, the program switches to Bing-based search. It still works reliably, ensuring the system stays functional even when external APIs fail.

---

## Smart Profile Matching

To match the correct LinkedIn profile, the app uses fuzzy matching and cosine similarity.

Fuzzy matching helps find names that are similar but not exact—like connecting “Jon” with “Jonathan.”

Cosine similarity looks at how similar two pieces of text are. For example, if two job descriptions are worded differently but mean the same thing, the app can still detect a match.

---

## **Income and Location Estimation**

Based on the job title found in the LinkedIn profile, the app estimates income and work location. If the result seems unrealistic—such as a salary over \$500K—it’s marked as “Unknown” to avoid skewing the data.

---

## **Clean, User-Friendly Interface**

The design is easy to use and clearly organized. Users can see search progress in real time, and they have the option to stop or restart a search at any point.

---

## **Exporting Results**

All results can be downloaded as a CSV file. This makes it easy to store, analyze, or share the data after a search is complete.

---

## **What Works Without SerpAPI**

Even without SerpAPI, the following still work:

- Bing-based search
  - Fuzzy and cosine similarity matching
  - Graduation year estimation
  - Full user interface and download options
- 

## **What Requires SerpAPI**

Some advanced features do need SerpAPI:

- Estimating location
  - Income estimation based on job title
  - Structured job info from Google search results
- 

## **Project Status**

This version of the app is fully stable, easy to use, and ready for testing. More updates are planned to further improve accuracy and add new features.

---