



JobScout

Mentors: Sohan Gharpure, David Mandelbaum

Group Members: Nessim Betesh, Owen Caplan, Adir Koenig, Sam Stolarov, Ilan Wurmman

GitHub Repository: <https://github.com/SM24-Industrial-Software-Dev/Job-Scout>

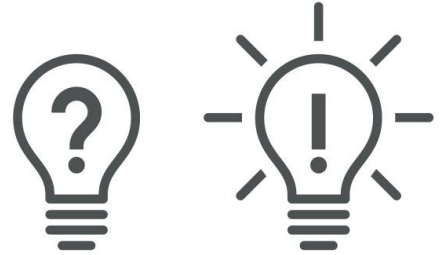
JobScout Website: <http://18.191.83.191:8501>



Introduction

Problem: There are many websites for job searching, and it takes a lot of time for users to find every relevant job as well as to stay up to date on new job postings.

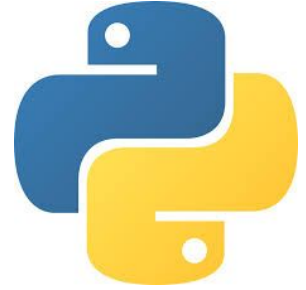
Solution: A website that aggregates jobs from various websites into one database and sends notifications with new job postings based on a user's job preferences.



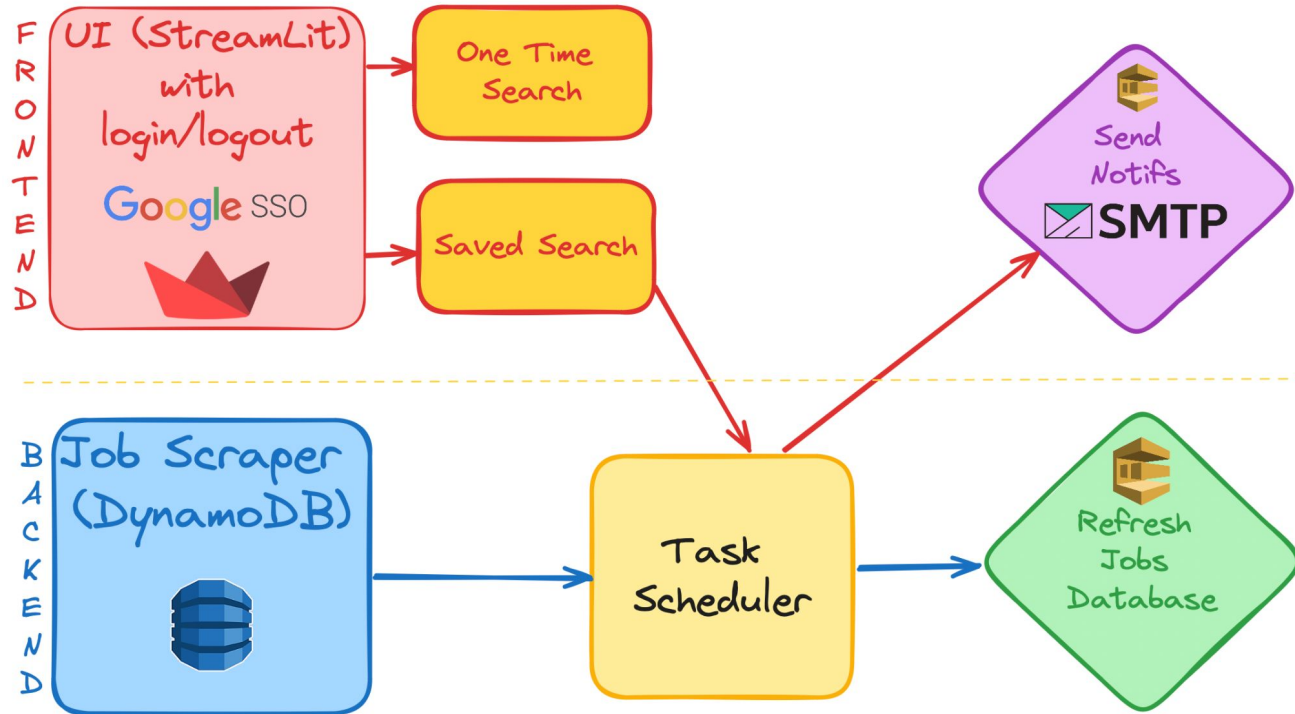


Tools and Frameworks:

- Python for backend code development
- Streamlit for UI
- FastAPI for UI endpoints
- AWS DynamoDB for job database and task scheduling
- Boto3 for Python interacting with DynamoDB
- AWS SQS for execution services
- AWS EC2 for deployment
- Flask for Google login
- SMTP for email notification



Software Design



Job Scraper

- Created 2 working job scrapers:
 - LinkedIn (Primary Job Scraper)
 - JobSpy (LinkedIn, Indeed, ZipRecruiter, Glassdoor)
- Challenges:
 - Deploying the JobSpy scraper
 - Need for third-party subscriptions

Items returned (50)

<input type="checkbox"/>	job_id (Number)	company	description	link	location	title
<input type="checkbox"/>	7	Experian	N/A	https://ww...	United States	Senior Relationship Manager - U...
<input type="checkbox"/>	47	ClearGov	N/A	https://ww...	United States	Manager of Client Success
<input type="checkbox"/>	8	Schuman C...	N/A	https://ww...	Bakersfield,...	Regional Sales Manager
<input type="checkbox"/>	32	GRAND Me...	N/A	https://ww...	Pawhuska, OK	General Application
<input type="checkbox"/>	44	JENNINGS ...	N/A	https://ww...	North Vern...	BCE Teacher
<input type="checkbox"/>	39	KENTECH	N/A	https://ww...	Chicago, IL	LMS MANAGER CONTRACTOR
<input type="checkbox"/>	10	Mondo	N/A	https://ww...	New York Cl...	Junior Business Analyst
<input type="checkbox"/>	31	Actalent	N/A	https://ww...	Albuquerque...	Cad Designer
<input type="checkbox"/>	3	UCLA Health	N/A	https://ww...	Los Angeles...	Patient Care Coordinator/Surger...
<input type="checkbox"/>	38	Actalent	N/A	https://ww...	Chicago, IL	Architectural Designer
<input type="checkbox"/>	2	SoFi	N/A	https://ww...	United States	Workplace Coordinator
<input type="checkbox"/>	21	Discovery ...	N/A	https://ww...	Annapolis, ...	RN Registered Nurse
<input type="checkbox"/>	40	MP: Wired f...	N/A	https://ww...	Clarksville, IN	Lab Assistant
<input type="checkbox"/>	14	12 Oaks Se...	N/A	https://ww...	Houston, TX	Caregiver
<input type="checkbox"/>	18	12 Oaks Se...	N/A	https://ww...	Houston, TX	Caregiver

Items returned (130)

<input type="checkbox"/>	job_id (Number)	company	link	location	title
<input type="checkbox"/>	10	Therap...	https://www.glass...	null	MD Psychia...
<input type="checkbox"/>	82		https://www.linke...	Seguin, TX	Part-Time S...
<input type="checkbox"/>	31	Descript...	https://www.indee...	Glendale, C...	Medical Ass...
<input type="checkbox"/>	58		https://www.linke...	New York, NY	Creative Dir...
<input type="checkbox"/>	3	searchi...	https://www.glass...	null	Office Assis...
<input type="checkbox"/>	65	Securi...	https://www.ziprec...	Chicago, IL,...	Armed Secu...
<input type="checkbox"/>	38	Descript...	https://www.indee...	Salt Lake Ci...	Registered ...
<input type="checkbox"/>	96		https://www.linke...	Santa Barb...	Locum Ten...
<input type="checkbox"/>	123		https://www.linke...	United States	Virtual Rese...
<input type="checkbox"/>	2	mary:**...	https://www.glass...	null	Business De...
<input type="checkbox"/>	120		https://www.linke...	Hampton, GA	Training an...
<input type="checkbox"/>	21	Descript...	https://www.indee...	Provo, UT, US	Phlebotomist
<input type="checkbox"/>	40	Descript...	https://www.indee...	South Jord...	Registered ...



Distributed Task Scheduler

- Handles two types of tasks:
 - Notifications
 - Refresh
- Checks every minute for executable tasks and executes them by sending Notifications to Notif SQS Queue and sending Refresh to Refresh SQS Queue.
- Utilizes multiple executor objects in order to distribute executions

Tasks Table:

<input type="checkbox"/>	task_id... ▾	company ▾	created ▾	interval ▾	location ▾	title ▾	type ▾	user_id ▾
<input type="checkbox"/>	0		1722380640	PT6H			refresh	
<input type="checkbox"/>	2	Yeshivas Ra...	1722380640	P1D	New York	Talmid	notif	107130786...
<input type="checkbox"/>	1	Google	1722380640	P1D	New York	Software E...	notif	107130786...

Executions Table:

<input type="checkbox"/>	task_id (Number) ▾	segment (Number) ▾	next_exec_time
<input type="checkbox"/>	2	3	1722467040
<input type="checkbox"/>	1	2	1722467040
<input type="checkbox"/>	0	1	1722402240



Here are the job search results for in at UCLA Health:

[[{"title": "Patient Care Coordinator/Surgery Scheduler - Neurosurgery", "company": "UCLA Health", "location": "Los Angeles, CA", "link": "https://www.linkedin.com/jobs/view/3988478932"}, {"title": "Patient Access Representative - Westwood (Part-Time)", "company": "UCLA Health", "location": "Los Angeles, CA", "link": "https://www.linkedin.com/jobs/view/3988480444"}, {"title": "Patient Service Rep - East-West Medicine, Westlake Village", "company": "UCLA Health", "location": "Westlake Village, CA", "link": "https://www.linkedin.com/jobs/view/3988477768"}, {"title": "Clinical Research Coordinator - Human Genetics", "company": "UCLA Health", "location": "Los Angeles, CA", "link": "https://www.linkedin.com/jobs/view/3988479493"}]]

↶ Reply

↷ Forward

SQS Listeners and Email Service

Refresh Listener:

- SQS Listener that runs the scraper at an interval decided when booting up the system, and refreshes our jobs database.

Notification Listener:

- SQS Listener that listens for executable tasks each minute and then passes the saved search preferences (such as location, company, job title) and User ID to our email service.
- Created a notification sender with Gmail SMTP that retrieves the user's email address that they used when logging in and sends the results to that address.

UI and Login

- Created a straightforward user interface with Streamlit:
 - Includes a page where logged-out users can execute instant searches
 - Includes a pages where logged-in where users can execute AND save searches
- Used Flask for login-logout logic:
 - The app includes a login button that redirects to a Google login page and a logout button to return to the standard page.
 - Logged-in users are assigned an ID which allows them to save searches

Menu
Login with Google

JobScout

JobScout is a web application that simplifies job searching by querying multiple job listing sites and saving user-defined searches. Utilizing a distributed search schedule, it regularly updates the job listing database and notifies users of new opportunities. The application also features a user-friendly web interface for manual queries and real-time results.

Job Title:

Location:

Company:

How often would you like to be notified?:

Search

Sign in with Google

Choose an account

to continue to
ec2-18-191-83-191.us-east-2.compute.amazonaws.com

1 Owen Caplan
ocaplan@mail.yu.edu

2 Use another account

To continue, Google will share your name, email address, language preference, and profile picture with ec2-18-191-83-191.us-east-2.compute.amazonaws.com.

English (United States) Help Privacy Terms

Logout

JobScout

JobScout is a web application that simplifies job searching by querying multiple job listing sites and saving user-defined searches. Utilizing a distributed search schedule, it regularly updates the job listing database and notifies users of new opportunities. The application also features a user-friendly web interface for manual queries and real-time results.

Please notify us if you received any security warnings when using JobScout.
If you received a security warning, you may need to terminate after 24 hours.
Refer to our [FAQ page](#) for more information.

Job Title:

Location:

Company:

How often would you like to be notified?:

Search

My Searches

Display My Searches

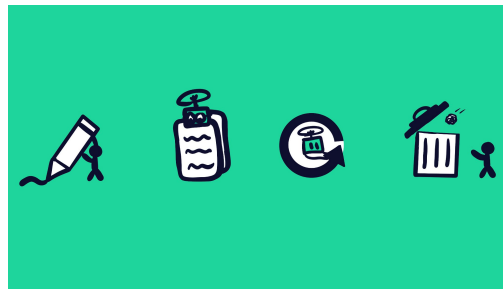
Software Development Skills Gained:

- Git/GitHub
- Stand-up meetings
- Working on a project as a team
- Local to Deployment
- AWS - IAM, EC2, SQS, DynamoDB



What's next

- Load Balancer
- Deploy the newer job scraper
- Improve website and email formatting
- Search suggestions
- More jobs scraped
- Delete/update searches





Demo

<https://youtu.be/f4ell1FLbGo>