Matthew Roseman

Education

Case Western Reserve University
Cleveland, OH | May 2017
BA in Computer Science
Minor in Mathematics

Experience

Hubspot
May 2021 - Nov 2022 | Backend Software Engineer | Remote
Java, SQL, Jersey, Guice (dependency injection, i.e. Spring alternative)

I worked on the Content Setup team, which was responsible for the transferring of user's blogs, knowledge bases, and site pages from their old sites into their HubSpot CMS..

- **Form Extractor**: Built extractor to scrape forms off customer's site and move to their HubSpot CMS
    - Wrote a heuristic algorithm able to accurately extract the human-readable labels in forms
    - Robust enough to work on all major third-party form builders
- **Cloudflare Allowlist**: Caught major bug blocking us from importing user's content from a popular competitor's site (**~70% of the imports** for knowledge bases)
    - Determined cause of outage was Cloudflare, and I coordinated communication with competitor, infrastructure teams, and internal contacts with Cloudflare to resolve the issue
- **Web Crawler**: Maintained web crawler built on Internet Archive's Heretrix crawler, used throughout HubSpot
- **Debugging**: Debugging system primarily responsible for orchestrating customer's import of content to **99.99% availability**

- **Work Environment**: Worked remotely in a strongly asynchronous work environment: involved being proactively communicative while taking ownership of the tasks assigned to me

Find.Jobs
May 2018 - April 2020 | Full Stack Software Engineer | Cleveland, OH & Remote
Python, Django, Redis, Elasticsearch, SQL, RabbitMQ

This was a small startup of 10-12 employees. The product was a job board website which had a unique website and domain for every job category, e.g., nursing.jobs, engineering.jobs, ohio.jobs. There was a single pool of jobs, and each website had its own filter.

- **Job Import**: Built job import system to handle **~2-3 million** jobs daily in **~1 hour**
  - Included downloading, normalizing multiple data formats, inferring missing fields, and storing jobs in SQL and Elasticsearch
- **Job Filtering**: Created a filtering system to the job boards search
  - Feature parity with Indeed's job filtering
  - Involved optimized Elasticsearch queries hitting **~10-20 million jobs** in less than a **2-3 seconds**, even in deep pagination
  - Built to be optimized for SEO discoverability
  - Same filtering system was used to dynamically build millions of sitemaps
- **Web Crawler**: Built web crawler tailored to premium customers' smaller job boards
  - Built with Python Scrapy
  - Built to be configurable to the point that adding new customer's websites took **~1 day of work** with no code changes

Keyfactor
May 2015–Aug 2015, May 2016–Aug 2016 | SE Intern | Independence, OH
ASP.NET, C#, SQL

Medium sized company of ~50 people. Product was an app that helped large companies manage their own site's certificates used for security purposes.

- Worked on general web development features on the frontend and backend, including modifications to a large SQL database
- Worked in ASP.NET with C#

● Work also involved some cursory Cryptography knowledge

Skills
Languages
● Python (6 years)
● Javascript (6 years)
  ○ ES6+
  ○ Node.js
● Java (2 years)
● Ruby
● Go

Databases
● SQL
● Elasticsearch
● MongoDB
● Redis

Frameworks/Libraries
● Django (3 years)
● Flask (2 years)
● Scrapy
● Heretrix
● Rabbitpy/Pika
● SQLAlchemy
● React (2 years)
● JQuery
● Bootstrap
● Selenium
● Webpack
● Babel

Infra
● Docker (2 years)
● GCP (2 years)
● Kubernetes