

## SKILLS

**LANGUAGES:** Python, C++, C, Groovy, Bash

**TECHNOLOGIES:** Docker, Jenkins, Git, AWS, Flask, SQL, RESTful Services, Microsoft SQL Server, Pandas, NumPy, Jira, Agile Development

## EMPLOYMENT

### SOFTWARE ENGINEERING FOR DATA SCIENCE

Jan. 2020 to Apr. 2020

Thomson Reuters

- Designed and developed a REST API for an NLP model that is currently used by 6 internal teams using Flask, Docker, SQLite and AWS.
- Optimized and parallelized scripts that run the Flair NLP model to make them 350% faster and saved over 400 hours of processing time.
- Built a pipeline to extract 25 million rows of legal data from a Microsoft SQL Server, run Flair and Spacy NLP models to remove personal information and then analyze them using Pandas.

### UNDERGRADUATE RESEARCHER- SOFTWARE ENGINEERING

Sep. 2019 to Dec. 2019, Jan. 2019 to Apr. 2019

SWAG Lab, University of Waterloo

- Collaborated with Prof. RG. Kula at NAIST, Japan on an empirical study of commit histories in open source projects.
- Created scripts in Python and Bash to crawl GitHub and mine the commit histories of 30,000+ users and to identify patterns in commit behaviour such as frequency, nature of commit, quality, etc.
- Classified commits based on the nature of change and analyzed data to find meaningful conclusions using Pandas.

Published initial findings as sole author at The International Conference of Software Engineering 2020: [bit.ly/3a8iV4w](https://bit.ly/3a8iV4w)

Submitted further findings as main author to IEEE Software Journal: [bit.ly/2Z1XpvJ](https://bit.ly/2Z1XpvJ)

### BUILD ENGINEER - DEVOPS

May 2019 to Aug. 2019

Sandvine Inc.

- Built a pre-merge CI pipeline for 4 repositories to run a battery of tests affecting 50+ commits/week using Groovy in Jenkins.
- Incorporated Docker-compose in the build system and reduced the number of instructions required to build services containing multiple docker images.
- Replaced NFS with SSH on build scripts for file transfer and thereby decreased copying times by 20% and stopped large builds from timing out.

### ROBOTICS INTERN

May 2018 to Aug. 2018

Trilogy Technologies

- Programmed a UR5 robotic arm and a Robotiq gripper to retrieve and return books in a library by generating MODBUS instructions in real-time.
- Programmed a motorized platform to make the arm autonomous using OpenCV and a network of Raspberry-Pis.
- Implemented an RFID system for the robot to identify unique books and developed an algorithm to identify the location of books with respect to each other.

## PROJECTS

### 'COMPLETE' - OVERALL WINNER AT HACK THE NORTH 2019 (1500+ PARTICIPANTS)

- Built a VSCode extension that produces relevant code snippets by understanding the meaning of docstrings the user writes and searching for code in open-source repositories.
- Created the backend by scraping GitHub's Semantic Search engine, using Difflib to filter search results and TF-IDF and regex to filter parameters.

### SEGARAGE.ORG - PUBLISHED IN ACM SIGSOFT

- Worked in a team of 3 developers to develop a platform that archives tools developed through software engineering research using AWS, Elasticsearch, MySQL and Flask.
- Represented this project at the International Conference of Software Engineering (ICSE) 2019 and processed over 80 tools from ICSE conferences.
- Published report on SEGArage: [bit.ly/3a3sUYb](https://bit.ly/3a3sUYb)

## EDUCATION

University of Waterloo

2018 to 2023

Software Engineering, Honours

Selected Coursework: Programming Principles (C), Data Abstraction and Implementation (C++, OOP), Software Engineering Principles, Data Structures and Data management, Introduction to Database management, Sequential Programs (Compilers).