

# bill sheng

billsheng.com  
github.com/billxsheng  
linkedin.com/in/billxsheng  
bxsheng@uwaterloo.ca

## skills

Languages: Java • Python • Ruby • SQL • JavaScript • TypeScript  
Frameworks: Eclipse Vert.x (Reactive) • Ruby on Rails • Node • React  
Data: Hadoop • Spark • Kafka • Cassandra • MongoDB • Elasticsearch • MySQL  
Infrastructure & Tools: Docker • GraphQL • Git • Jenkins • Azure

## experience

### Backend Engineer Intern @ Shopify

Toronto, Jan 2021 – Apr 2021

- Worked on Shopify's Google sales channel application, shipped backend features and optimizations with Rails to **400,000+** merchants
- Developed onboarding flow for latest Google sales programs with Ruby, GraphQL, and Google API's, launched beta to **100+** shops
- Refactored current onboarding data model using Rails AR/LHM migrations, simplifying the addition of new programs/analytics events
- Optimized Redis infrastructure by removing unused shards and updating clock job queuing logic, cutting average RAM usage by **~15%**

### Software Engineer Intern @ NVIDIA

Santa Clara, May 2020 – Aug 2020

- Led the end-to-end development of a reactive ETL pipeline (5M+ msgs/day) that processes critical test data from **100+** NVIDIA projects
- Wrote a Python script that purges Artifactory artifacts based on data retention rules, removing **~5GB** of unused resources every month
- Converted a monolithic app into a reactive Java microservice using Vert.x, raising throughput by **~6x** and enabling dynamic scalability
- Leveraged ApacheBench and JProfiler to identify bottleneck operations in posting API, refactored code to decrease latency by **~23%**

### Full Stack Developer Intern @ LCBO|next

Waterloo, Sep 2019 – Dec 2019

- Built a cron-based web scraper that extracts marketable content from **950+** webpages, increasing app conversion paths from **9 to 400**
- Developed a Node ETL pipeline that tags products with new properties to improve search precision, decreasing search exits by **25%**
- Designed an indexing strategy with Elasticsearch aliases that prevents data loss during API crashes in **3** large-scale data pipelines
- Co-built and deployed a Node recommendation engine, leveraging compound Elasticsearch queries to address customer indecision

### Software Developer Intern @ OpenText

Toronto, Jan 2019 - Apr 2019

- Developed a monitoring service with Spring and Thymeleaf that reports on the state of Java Agents, notifying users of build timeouts
- Refactored a high-profile deprecated dashboard into an Angular PWA used by **800+** stakeholders at annual enterprise conference

### Infrastructure Specialist @ Canadian Tire Corporation

Toronto, May 2018 – Aug 2018

- Streamlined inefficient VM configuration processes by writing PowerShell automation scripts used by **50+** employees across IT teams

## leadership

### Software Developer @ UWBlueprint

Waterloo, Sep 2019 – Dec 2019

- Built a web platform for a non-profit accelerator group (**~100** users) aimed to scale and manage their application selection process
- Co-designed a document data model to represent application flows and developed CRUD operations with React/Redux and MongoDB

## projects

### Oubre Sentiment Analysis Platform (Java, Kafka, Spark, Cassandra)

[Repo](#) [Slidedeck](#)

- A data platform that performs sentiment analysis on live tweets and writes periodic MapReduce aggregations into a data warehouse

### WeWatch (React, Node, Redis)

[Repo](#)

- A swipe-based web application leveraging web sockets (Socket.IO) and a Redis store to match groups with movies of mutual interest

### NBA Draft Analysis (Python, Scikit-learn)

[Repo](#) [Case Study](#)

- Implemented k-means clustering algorithms with Jupyter Notebook to study playstyle trends in NBA draftees over the past 30 years

## education

### University of Waterloo | 4A Management Engineering (CGPA: 3.7)

Sep 2017 - Apr 2022

Relevant Courses: Database Systems, Data Mining, Data Structures and Algorithms, Object-Oriented Programming, Search Engines