Matt D'Souza

matt.dsouza@edu.uwaterloo.ca
mattdsouza.com
DSouzaM

Skills

- Fluent in Java, Python, C++, Scala, Ruby, OCaml, SQL
- Data engineering experience using Hive, Presto, MongoDB, Redis, Hadoop, and Spark
- Infrastructure experience using AWS, Chef, Prometheus and Google Analytics

Experience

May–Aug 2018 **Software Engineering Intern** | *Snowflake Computing* | San Mateo, California

- Optimized SQL client's handling of array binds to protect the database from degraded performance
- Implemented compiler and client support for multi-statement query execution
- Created a UDF to automatically parse and infer the file format and schema of CSV files
- Improved SQL client telemetry to support real-time client monitoring and troubleshooting

Sep–Dec 2017 **Relevancy Engineering Intern** | *Wish* | Toronto, Ontario

- Reduced the time to build product image similarity graph from 20 hours to 30 minutes
- Developed a versioning scheme for Redis to enable consistent data updates across all Redis nodes
- Implemented backup, restore, and monitoring of graph databases using AWS and Prometheus
- Created a system to monitor changes in the performance of machine learning models

Jan–Apr 2017 **Data Platform Engineering Intern** | *Shopify* | Toronto, Ontario

- Built and maintained connectors to extract data from Google Cloud and other third-party APIs
- Reduced overlap between incremental extraction jobs by 95% after using Spark to investigate late-arriving data from historical logs
- Automated environment setup and project build process to improve developer efficiency

May–Aug 2016 **Software Engineering Intern** | *Veeva Systems* | Toronto, Ontario

- Built a Spring service to asynchronously send backend metrics to Google Analytics
- Spearheaded and took ownership of rewriting of the application's breadcrumb navigation system

Projects

2018 **Tiger Compiler** | DSouzaM/OCaml-Tiger
OCaml implementation of the Tiger language

- Used ocamllex and ocamlyacc to create a lexical analyzer and parser for the front-end
- Implemented a type-checker with support for recursive and mutually-recursive data types

2017 **Language Analysis** | DSouzaM/LanguageAnalysis
Comparison of GitHub projects by programming language

- Implemented tokenization and keyword extraction algorithms to generate features from READMEs
- Experimented with clustering to group repositories and look for programming language trends

Education

2015–2020 **Candidate for Bachelor of Software Engineering, 4.0 GPA** | *University of Waterloo*

- Undergraduate Research Assistant (2016): Experimented with Bazel and Java 9 compatibility for an annotation-based type-checking framework

Interests

- Court & beach volleyball
- Compilers, programming languages
- Acoustic guitar, electric keyboard