James Liang

+1-514-779-6503 | james.liang.cje@gmail.com | linkedin.com/in/jamesyueliang | github.com/pooky1955

Education

McGill University — Montreal, Canada
Bachelor of Science (BSc) Honours Mathematics and Computer Science — *August 2023 – May 2026*

Marianopolis College — Montreal, Canada
Diploma of College Studies (DCS) Honours Pure and Applied Sciences — *August 2021 – June 2023*
Activities: President of MariHacks, Founding President of the AI Club, Co-President of the Mari Math Club

Hosted MariHacks 2023, Montreal's largest CEGEP hackathon, for the first time in-person at Marianopolis College, raised 4.5k+$ through 7 sponsors and welcomed 178 participants for 2 days. Taught web development, data structures & algorithms, math competitions, and artificial intelligence workshops every week to 25+ attendees.

Technical Skills

Languages: Python, JavaScript / Typescript, HTML, CSS, Bash, Java, Go, SQL, R, C, C++
Developer Tools: Linux / Unix, Git, Docker, Nginx, Jupyter, Maven, Gradle, Node.js, Webpack, Travis CI, Excel
Libraries/Frameworks: (Frontend): React, Redux, Gatsby (Backend): Flask, Spring Boot, Node.js HTTP, Express

Experience

Full-Stack Developer and Data Scientist — August 2021 – Present
University of Montreal - LabTNS | NLP, React, Python, Flask, Java, Spring Boot, Docker, Nginx — *Montreal, Canada*
- Developed a full-stack web application using Flask, React, Java, Spring Boot, and GraphDB to retrieve more than 25k+ standardized clinical concepts under 500ms such as medication component, form, and route of administration
- Implemented a secure Spring Boot backend API using Bearer authentication, documented with Swagger
- Maintained server infrastructure using Nginx load balancing and handled troubleshooting of 3 servers for 2+ years
- Aligned 4+ national drug databases to OCRx's ontology totalling 76k+ entities using RegEx and NLP heuristics

Data Science Intern — May 2021 – August 2021
Boehringer Ingelheim | Python, R, Pandas, Scikit-Learn — *Burlington, Canada*
- Preprocessed 15+ years of physician history using R and Pandas to optimize email click-through-rates
- Engineered a data-driven solution using machine learning techniques such as XGBoost, LDA, and SVM to predict optimal email send time for physicians, achieving a top 3 accuracy rate of 80% (2.6 times higher than baseline)

Biostatistics and Data Science Intern — May 2020 – August 2020
Boehringer Ingelheim | R, RShiny, Python, Pandas, PostgreSQL, Travis CI — *Burlington, Canada*
- Programmed an RShiny app that allows users to create complex SQL-like queries through a simple interface to query 211k+ clinical trials in under 2 seconds from clinicaltrials.gov
- Leveraged NLP and Google's Tesseract OCR to extract consent indicators from 17+ years of informed consents
- Streamlined software production process by implementing Travis CI for continuous integration and Pytest for unit testing for 500+ lines of code
- Improved Synthea, an open-source synthetic patient generator, by generating localized (Canadian) demographics data for 1854 hospitals and 967 counties (using Statistics Canada open datasets)

Software Projects

Inspiratiq | *React, Gatsby, Material UI, GraphQL, Stripe, Netlify functions* — January 2020
- Developed an end-to-end serverless web application serving over 100+ customers on Netlify for Inspiratiq's artwork using Gatsby, Material UI, and GraphQL
- Integrated payment within the website using Stripe Checkout Sessions through Netlify functions for 300+ orders

iNews | *React, Material UI, Flask, Keras* — January 2020
- Constructed a React frontend application that detects fake news by performing validation against 30+ articles from Google News, fact-checking websites and source credibility verifiers
- Performed stance detection on Google News article through a custom multi-channel convolutional neural network on Word2Vec embeddings to achieve an 85% accuracy vs 82% from state-of-the-art

SmartCT (1st Place at PharmaHacks) | *React, Sklearn, Flask, BeautifulSoup* — November 2019
- Created a web application using React and Material UI to predict the success of clinical trials in under 24h
- Performed automatic asynchronous scraping of 100+ clinical trials under 5 seconds using BeautifulSoup and Requests-HTML for any given trial