

# Kalvin Khuu

**LinkedIn:** <https://www.linkedin.com/in/kalvin-khuu/>

**GitHub:** <https://github.com/kalvinkhuu>

**Languages:** French (Native), English (Native), Vietnamese (Intermediate)

**Email:** [kalvin.khuu@gmail.com](mailto:kalvin.khuu@gmail.com)

**Mobile:** (514) 992-7442

## EDUCATION

### Ph.D. in Software Engineering

McMaster University

Sep 2024 – Present

Hamilton, ON

- **GPA:** 3.9/4.0
- **Topics:** AI safety, Assurance Cases, MLOps
- Supervised by [Dr. Sébastien Mosser](#) under the [FATES-MLOps](#) project

### Bachelor of Eng. in Software Engineering

McGill University

Sep 2020 – May 2024

Montreal, QC

## TECHNICAL SKILLS

**Languages:** Python, Java, TypeScript, JavaScript, C, C#, C++, HTML/CSS, Unix/Bash

**Databases:** SQL, Oracle

**Frameworks:** Django, React, Angular, Node.js, Spring, TensorFlow, RabbitMQ, jQuery, Cucumber

**Cloud & DevOps:** AWS, Azure, Docker, Git, GitHub, Maven, Gradle

**Tools:** IntelliJ, VS Code, Postman, Jira, Figma, Confluence, BitBucket, Android Studio

## PUBLICATIONS

### MODELS | *When Safety Claims Fail Silently: A Modelling Approach to AI Assurance*

Oct 2026

- Proposed a modelling language for AI safety claims that turns natural-language model and system cards into structured justification diagrams, exposing where safety arguments break down across the AI lifecycle.

### MODELS | *Modelling Justifications to Support and Operationalize Machine Learning Pipelines*

Oct 2026

- Co-authored a tutorial where we modelled and validated ML pipeline design rationale using jPipe to combat "decision drift", integrating automated conformance checks into CI/CD for reproducible, requirement-traceable pipelines.

### MoDRE | *Safety First! Modelling Requirements from GPT-5 System Card using Lightweight Safety Models*

Aug 2026

- Lead-authored a paper where we performed an in-depth manual analysis of the 60-page GPT-5 system card, eliciting 74 claims and 22 implicit safety requirements, classified against the MIT AI Risk Repository taxonomy.
- Analysis reveals uneven, tightly coupled, and under-specified requirement coverage across privacy, safety, and misinformation domains.
- Applied justification diagrams to formally model requirement compliance (e.g., chemical/biological capability thresholds), surfacing OpenAI's heavy reliance on safeguarding, benchmarks, and external red-teaming as evidence.

### RE | *Validating Requirements of Machine Learning Pipelines with Justification Models*

Aug 2026

- Co-authored a tutorial where we modelled and validated ML pipeline design rationale using jPipe to combat "decision drift", integrating automated conformance checks into CI/CD for reproducible, requirement-traceable pipelines.

### ECMFA | *Capturing and Organizing Reusable Interaction Practices Using Justification and Feature Models*

Jun 2026

- Developed a novel framework combining safety-inspired justification diagrams with feature-model variability management to make tacit best-practice knowledge explicit, composable, and reusable.
- Demonstrated automated composition (assembly and refinement operators) of individual practices into coherent, auditable argumentation structures using the JPIPE and UVL/FLAMAPY toolchains.

### SCC-BOOK'26 | *Role of MLOps to validate the Requirements of AI Models using Lightweight Assurance Cases*

May 2026

- Proposed a lightweight assurance-case approach integrating MLOps practices to support requirements validation and certification of AI systems within a book chapter.

### ICSE-NIER | *Model Cards for Responsible AI: Stop Carding, Start Modelling*

Apr 2026

- Lead-authored a paper that evaluates current AI/ML documentation practices and introduces an actionable, verifiable alternative to narrative model cards.
- Extracted and categorized **32 recurring safety patterns** from provider model cards; quantified **inter-provider inconsistency** (only ~22% shared coverage across patterns).
- Operationalized model-card content into structured assurance cases for **CBRN** and **cybersecurity** risks using evidence from **GPT-OSS**, **Claude 3**, and **Gemma 3n**.

## WORK EXPERIENCES

---

### McMaster University

Sep 2025 – Present

*Graduate Lead Teaching Assistant*

*Hamilton, ON*

- Lead Teaching Assistant for SFWRENG/COMPSCI 3RA3 and SFWRENG 3A04, providing academic support through office hours, tutorials, and student mentoring.
- Co-designed course materials, assignments, and assessments in collaboration with the instructional team to improve learning outcomes.

### McSCert

May 2025 – Aug 2025

*Research Assistant III*

*Hamilton, ON*

- Worked on jPipe, an open-source justification modeling language to support software maintenance activities
  - \* **Research:** Designed and developed composition operators within jPipe, allowing for reusability and scalability
  - \* **ML Fairness Techniques:** Implemented justification frameworks related to fairness definitions and **FATES** principles
  - \* **Academic Presentation:** Presented the system prototype in academic settings to gather user feedback and demonstrate its role in supporting fairness-aware ML development

### Desjardins Group

May 2023 – Aug 2023

*Software Analyst Intern*

*Montreal, QC*

- Worked with the team of TFE, *Electronic Fund Transfers*, to modernize the current application of DRD, *Direct Deposit-Withdrawal*
  - \* **Quality Improvement:** Successfully solved character encoding discrepancies to achieve 100% precision per 1000 files
  - \* **Query Design:** Established and optimized **SQL/Oracle** queries to retrieve files exhibiting problematic behavior, allowing for targeted analysis and troubleshooting of identified issues
  - \* **Validation:** Implemented tests to validate the effectiveness of the solutions for 4 different character encoding

### The Quebec Ministry of Public Security

May 2022 – Mar 2023

*Frontend Developer Intern*

*Montreal, QC*

- Worked on IGO2, a free open source Web Geospatial solution used by various government departments
  - \* **User Customization:** Revamped the **UI/UX** to enable user customization, resulting in a 25% reduction in clutter
  - \* **Interactive OpenLayers Overlay:** Enabled users to interact with geographical data effectively by creating 7 new features
  - \* **User-Centric Design:** Prioritizing usability testing and feedback loops to align the final product with user expectations

## SOFTWARE DESIGN PROJECTS

---

CogSurf | *JavaScript, HTML/CSS, Bootstrap, jQuery, Vercel*

Aug 2024 – Aug 2025

- Designed, implemented and deployed the website for the Cognitive Science Undergraduate Research Forum

KnowledgeKnot (ML Engineer) | *Django (Python), Angular (TypeScript), Docker, OpenAI*

Sep 2023 – Apr 2024

- Engineered a **NLP** model with **BERT** to classify & cluster info, and displaying it on an interactive web interface