

# Zikun Fu

ML/NLP Research Engineer, Applied AI

📍 Oshawa, ON, Canada | 📞 (437) 922-1214 | ✉ [Zikun.Fu@ontariotechu.net](mailto:Zikun.Fu@ontariotechu.net) | 🌐 [zikunfu.github.io](https://zikunfu.github.io) |

🌐 [linkedin.com/in/zikun-fu](https://linkedin.com/in/zikun-fu) | 🐙 [github.com/ZikunFu](https://github.com/ZikunFu)

## Summary

MSc Computer Science graduate (Ontario Tech, 2026) working on language model systems grounded in structured data. My thesis formalized Database Entity Recognition and reached 93.2%  $F_1$  with a T5 + MiniLM pipeline, published at IEEE IRI 2025 alongside a public CC BY 4.0 benchmark dataset. I also build applied AI systems and evaluate them, including a SQL analytics assistant with measured grounding and abstention behaviour and automated pipelines gated on cost and human review. My methods work covers dataset curation and annotation design, benchmark construction, reproducible evaluation and failure analysis, and statistical comparison.

## Education

**Ontario Tech University**, MSc, Computer Science Jan 2024 – Apr 2026 | GPA: 4.24/4.3  
**Ontario Tech University**, BSc (Hons) Computer Science (Data Science) 2019 – 2023 | GPA: 3.86/4.3

## Research

- 🎓 **MSc Thesis:** [Database Entity Recognition using Language Models](#) Defended Apr 2026
- Formalized DB-ER as sequence labeling over {O, Table, Column, Value}. Co-created a 1,000-example human benchmark (Spider/BIRD) and a 15,026-example synthetic annotation pipeline via SQL AST parsing and integer linear programming. T5-Large reaches 80.8%  $F_1$ , and closed-world MiniLM verification raises it to 93.2%.
  - Public release: 🐙 [ZikunFu/DBER](#) (MIT) and 🗄 [Voice49/dber](#), holding 16,126 rows across four splits under CC BY 4.0.

## Publications

- Z. Fu**, C. Yang, K. Davoudi, K. Q. Pu. “Database Entity Recognition with Data Augmentation and Deep Learning.” *IEEE IRI 2025*, San Jose, CA. 📄 DOI: [10.1109/IRI66576.2025.00071](https://doi.org/10.1109/IRI66576.2025.00071)
- Ontario Database Day 2024: 📄 [Transforming Text-to-SQL Datasets into Closed-Domain NER Benchmarks](#)

## Experience

**Research Assistant** May 2026 – Present  
*Ontario Tech University*

- Building grounded LLM systems over structured business data, as sole developer on both. The first is an analytics assistant using DuckDB, a sqlglot SQL allowlist, and per-answer SQL provenance; against a held-out answer key it answered every supported question correctly and declined the rest rather than inventing figures. The second is an automated web-presence audit pipeline with cost and human-review gates.

**Teaching Assistant** Jan 2024 – Apr 2026  
*Ontario Tech University*

- TA every term from the start of my MSc through graduation, across compilers (Kotlin), scientific data analysis, programming workshops, and introductory CS and programming courses: lab support, grading, and exam marking.
- Built a LLM-assisted marking pipeline for scanned handwritten assignments (Claude via OpenRouter). It graded each question image against the model solution, raised confidence and legibility flags, and kept an audit trail, with the human marker re-scoring flagged items. Model cost was about \$0.15 per student.

**GREx Redevelopment Project Manager** Dec 2024 – Apr 2025  
*Frazer Faculty of Education, Ontario Tech University*

- Led the migration off a legacy Symfony/Laravel stack onto Dockerized LimeSurvey 6.x with MySQL, writing Python ETL against the REST API to import 3,400+ multilingual records across 60+ research groups.
- Built PHP plugin modules and Chart.js dashboards, deployed to production on DigitalOcean behind an Apache reverse proxy with automated TLS. I owned the migration plan, the data layer, and the cutover, and the platform is still running.

## Projects

---

### Hermes Clinic Analytics | [ZikunFu/hermes-clinic-analytics](https://github.com/ZikunFu/hermes-clinic-analytics)

2026

- Grounded AI analytics assistant for simulated clinic business data. A DuckDB engine restricted to reads sits behind a sqlglot SQL allowlist, with the presentation database kept physically separate from the hidden oracle tables; charts are built from query results and every answer reports the exact SQL behind it. On the published harness it matched the database on every answerable question and returned no invented values for the unanswerable ones, and offline and live bypass-safety suites verify the boundaries.

### Web Presence Report, automated local-SEO audit pipeline

2026 – ongoing

- Automated pipeline for a local optometry clinic turning Google Business Profile, reviews, PageSpeed, geo-grid rankings, and backlinks into a graded client-ready report with a 30/60/90-day action plan; cost gates before paid API calls and human-review gates for health-content compliance.

## Skills

---

**Languages:** Python, SQL, Kotlin, PHP, JavaScript, Bash

**ML/NLP:** PyTorch, Hugging Face Transformers (T5, BERT, MiniLM), scikit-learn; NER/sequence labeling, embeddings, semantic similarity, fine-tuning, evaluation harness design

**LLM Engineering:** grounded/tool-using agents, SQL validation gates (sqlglot), prompt design & caching, structured outputs, grounding/hallucination measurement, abstention, cost tracking; Claude API, OpenRouter, OpenAI-compatible APIs

**Data & Pipelines:** DuckDB, pandas, NumPy, ETL, data migration and validation, SQL AST parsing (sqlglot, pglast), constrained optimization (OR-Tools CP-SAT); MySQL, SQLite

**Backend/Infra:** REST API integration, Docker, Docker Compose, Apache reverse proxy, Linux, DigitalOcean, Git/GitHub, automated TLS (Let's Encrypt), pytest