

# Zikun Fu

📍 Oshawa, Ontario 📩 Zikun.Fu@ontariotechu.net 📞 (437) 922-1214 💬 LinkedIn 🐣 GitHub

## Summary

ML/NLP-focused Master's student building systems that connect natural language to structured data. Built a database entity recognition pipeline (outperforming widely adopted frameworks) with  $30\times$  synthetic data augmentation, and led a containerized platform migration (3,400+ multilingual records across 60+ research groups). Strong Python, PyTorch/Transformers, SQL, Docker, and REST API integration.

## Key Achievements

**NLP/ML Pipeline:** Built end-to-end Database entity recognition, outperform industry frameworks.

**Data Augmentation:** Generated synthetic data via novel pipeline; outperform human gold labels.

**Embeddings & Retrieval:** Designed novel DB grounding module; up to 10% performance gain.

**Full-Stack Migration:** Dockerized LimeSurvey 6.x (MySQL + Apache proxy); Python ETL via REST API importing 3,400+ multilingual records; DigitalOcean deployment with automated TLS.

## Skills

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

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

**Data & Pipelines:** pandas, NumPy, ETL, SQL parsing (pglast), constraint optimization (OR-Tools CP-SAT); SQLite, MySQL, PostgreSQL, MongoDB

**Backend/APIs:** REST API integration (LimeSurvey RC2, Google API), FastAPI (familiar), Docker, Docker Compose, Apache reverse proxy

**DevOps/Infra:** Linux, DigitalOcean, Git/GitHub, automated TLS (Let's Encrypt), cron, SSH

## Experience

### GREx Redevelopment Project Manager

Dec 2024 - May 2025

*The Educational Informatics Laboratory (EILab), OTU*

- Led full-stack migration to Docker (LimeSurvey + MySQL + Apache); Python ETL via REST API importing 3,400+ multilingual records across 60+ research groups.
- Built PHP backend modules and Chart.js dashboards for data visualization.

### Research Associate

Jan 2024 - Present

*Ontario Tech University*

- Collaborated with faculty in data Science research including DB-ER task formalization, data analysis, and deep learning model design.
- TA for undergraduate courses; integrated AI tools into teaching workflows.

## Research

### 2024 Fall - Present Master's Thesis: Database Entity Recognition for Text-to-SQL Systems

- DB-ER task formalization; created human and synthetic benchmarks; designed end-to-end DB-ER pipe.

## 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](#) ↗

## Education

**Ontario Tech University** — Master of Information Science

Exp. 2026 — GPA: 4.24/4.3

**Ontario Tech University** — BSc (Hons) Computer Science (Data Science)

2019 – 2023 — GPA: 3.86/4.3