Yuan Tian

tian211@purdue.edu | (+1) 765-426-6107 | yuan-tian.com

Education

Purdue University	*Aug. 2021 – Present*
Ph.D. (& M.S.) in Computer Science	GPA: 3.71 / 4.00
University at Albany, SUNY	*Aug. 2019 – May 2020*
B.S. in Computer Science	GPA: 3.97 / 4.00 (Dean's List)
Chongqing University of Posts and Telecommunications	*Aug. 2016 – Jul. 2019*
B.E. in Computer Science & Software Engineering	GPA: 3.84 / 4.00 (Ranked 1st)

Skills

- I am a full-stack developer capable of independently building user interfaces, servers with databases. I'm currently most comfortable with Python, and also proficient in Java, JS, SQL, and C/C++.
- I am an independent researcher capable of developing novel and effective ML, NLP, and HCI algorithms.

Work Experience

Applied Scientist / Machine Learning Engineer [Intern]   *May 2025 – Present*
Adobe, GenAI @ AEP | San Jose, CA, USA
- R&D at a novel semantic enrichment system for under-specified fields
- Robust richness scoring & clarification generation & interactive evaluation by populated queries
- To integrate with and benefit the Adobe XDM system

Research Assistant   *Aug. 2024 – May 2025*
Purdue University | West Lafayette, IN, USA
- Individual research (Papers accepted)
- Build Knowledge Graph for Software Supply Chain Security [NSF]

Applied Scientist / Machine Learning Engineer [Intern]   *May 2024 – Aug. 2024*
Adobe, GenAI @ AEP | San Jose, CA, USA
- R&D at a novel text-to-SQL domain adaptation system
- Data schema management & high-quality, customized text-to-SQL data generation
- Patented & Open-sourced & Accepted to IUI

Teaching Assistant   *Jan. 2024 – May 2024*
Purdue University | West Lafayette, IN, USA
- Deliver lectures on discrete mathematics, algorithms, and data structures (CS 182).

Research Assistant   *Aug. 2021 – Jan. 2024*
Purdue University | West Lafayette, IN, USA
- Individual research (Papers accepted)
- Build Knowledge Graph for Software Supply Chain Security [NSF]

Publications

EvoSchema: Towards Text-to-SQL Robustness Against Schema Evolution
[VLDB 2026] | Schema Evolution, Text-to-SQL, Benchmark, Perturbation Taxonomy
- Benchmark to test text-to-SQL robustness under schema evolution
- 10 Types of Schema Perturbations Based on a Hybrid Method (LLM + Heuristics)
- Fine-tuning LLMs on perturbed schemas benefits performance improvement

Selective Prompt Anchoring for Code Generation
[ICML 2025] | Attention Steering, Taylor Expansion, Logit Arithmetic, Attention Analysis, Code Generation
- Identified the attention dilution phenomenon as a root cause of code generation errors in LLMs.
- Proposed and mathematically proved a general attention-steering method for LLMs
- Integrate with Hugging Face API and support all Hugging Face LLMs

- Proposed an attention-based code generation pipeline, achieving new SOTA performance

Text-to-SQL Domain Adaptation via Human–LLM Collaborative Data Annotation
[IUI 2025] | Human-AI Collaboration, Domain Adaptation, Interactive Systems, Text-to-SQL, Data Augmentation
- Proposed a comprehensive framework (UI + backend) for schema editing, interactive text-to-SQL annotation, automated text-to-SQL data augmentation, and text-to-SQL dataset analysis
- Rigorous user study with 12 participants to evaluate usability and annotation efficiency

SQLucid: Grounding Natural Language Database Queries with Interactive Explanations
[UIST 2024] | Interactive Systems, Text-to-SQL, Database Interfaces, Grounding Theory
- Built a novel interactive SQL generation tool based on editable step-by-step explanations, visual grounding, and intermediate query executions
- Conducted two comprehensive user studies (30 participants) validating system effectiveness

Insights into NL Database Query Errors: From Attention Misalignment to User Strategies
[TiiS 2024] | Error Analysis, Attention Mechanisms, User Behavior
- Extended our previous text-to-SQL error analysis to include LLM analysis and attention studies.
- Demonstrate that models make errors when their attention does not align with human attention

Interactive SQL Generation via Editable Step-by-Step Explanation
[EMNLP 2023] | Grammar/Rule-based method, SQL Parsing, Explanations, Text-to-SQL, Clause Generation
- Proposed "editable step-by-step explanation", novel mechanism for SQL generation and repair
- Built a robust grammar-based SQL parser, a rule-based NL explanation generator, and a neural-symbolic clause-level SQL editing model for error correction

An Empirical Study of Model Errors and User Repair Strategies in NL-to-SQL
[IUI 2023] | Error Taxonomy, User Studies, Interactive Repair
- Developed a taxonomy of SQL errors produced by SOTA text-to-SQL models
- within-subjects study with 26 participants to evaluate three interactive systems and provide insights

Additional Projects

Supporting Construction Worker Well-Being with Multi-Agent Conversational AI
[CRC 2025] | Multi-Agent Systems, Conversational AI, Well-being
- Developed a conversational multi-agent system for construction workers' mental health support
- Agent customization with internal prompt generation and external RAG-based document upload
- A user study with 12 participants to demonstrate improved user engagement and support effectiveness through group chat with AI agents

Encountered-Type Haptic Display via Tracking-Calibrated Robot
VR, Robotics, Unity, Oculus, Haptics, Tactile Feedback, Calibration
- Developed a novel method ot simulate the haptic feedback in VR using the UR16e robotic arm
- Developed a robust tracking calibration algorithm for virtual-physical synchronization

Conversational Agent for SQL Generation
Conversational AI, Rule-based Clarification, Database Interfaces, UI
- Developed a conversational natural language interface for database query
- Developed a clarification method for disambiguation during generation and query repair after generation

Dancing Humanoid Robot
Robotics, Humanoid Robot, Calibration
- Designed the communication protocol between the upper computer and the ESP8266 chip inside the robot
- Enabled the robot to perform actions (e.g., walk) based on predefined parameters and calibrations.

Ultrasonic-powered Parking Management System
Automated Data Management, Ultrasonic Sensor
- Developed an automated parking management system based on ultrasonic sensors
- Link ultrasonic sensors, the Java virtual machine, and Microsoft SQL server through serial port communication.