

# Shenzhe (Cho) Zhu

7 Mirrow Court, Scarborough, ON, M1C 4Y7

+1 437-849-4752 | [cho.zhu@mail.utoronto.ca](mailto:cho.zhu@mail.utoronto.ca) | [linkedin.com/in/shenzhezhu0531](https://www.linkedin.com/in/shenzhezhu0531) [shenzhezhu.github.io](https://shenzhezhu.github.io)

## EDUCATION

### University of Toronto

Bachelor of Science in Computer Science

- Cumulative GPA: 3.97/4.0
- Honors&Awards: 2022-2023 Dean's List

Toronto, Canada  
Sept 2022 - Current

## RESEARCH INTERESTS

- Explainable AI (XAI), LLM Intepretability & Alignment, Trustworthy AI, LLM Agent

## RESEARCH EXPERIENCE

### PRADA Lab, King Abdullah University of Science and Technology

Collaborative Researcher

Remote  
June 2024 - Current

- **Synergizing Sparse Autoencoders and Representation Control for Personalization Interpretation**
  - Built experiment pipelines using sparse autoencoder (SAE) based on GPT-2 to analyze MBTI and Big Five corpus in batches aimed at identifying monosemantic features that represent personality traits.
  - Created feature activation table to link features with token activations in the corpus, constructed NoSQL database to enhance the efficiency of reading and writing the table.

### SocialAI Lab, University of Toronto

Research Assistant, Advisor: Prof. William Cunningham

Toronto, Canada  
May 2024 - Current

- **SC-Bench: Evaluating LLM-based Generative Agents through Social Cognitive Criteria**
  - Engaged extensive social cognitive criteria experiments for agents of 25+ pre-trained LLMs from 5 aspects to evaluate abilities and limitations of language models in social cognition and simulating human-like interactions.
  - Developed harmonized performance metrics for sub-experiments, designed automatic scoring mechanism based on the harmonic mean.
- **Minagen: A Minimal Testing Ground for Building Cognitive Architectures for Generative Agents**
  - Co-engineered cognitive and environmental simulation components to bridge the upstream Ollama and Langchain based LLM framework, with downstream cognitive architecture construction, enhancing the overall interoperability of the Minagen platform.
  - Enhanced the prompt standardization for agent-environment-cognition interaction, and developed comprehensive usage examples.

### CoNSens Lab, University of Toronto

Research Assistant, Advisor: Prof. Matthias Niemeier

Toronto, Canada  
May 2024 - Sept 2024(Expected)

- Evaluated emergent properties using neural network visualization tools for studying predictive coding and attention mechanisms
- Modified the structure of gradient-based visualization models such as LayerCAM and Grad-Cam++ to focus on single outputs.

### Department of Computer Science, University of Toronto

Machine Learning Researcher, Advisor: Prof. Xujie Si

Toronto, Canada  
Dec 2023 - Dec 2024(Expected)

- Crafted a survey paper as first author with analyzing 140+ papers on neural symbolic systems and deep learning interpretability.
- Curated and annotated the Yale Face dataset for facial feature analysis, and finetuned instance segmentation models like Mask-RCNN.

## WORK EXPERIENCE

### Urban Data Research Centre, University of Toronto School of Cities

NLP Analyst

Toronto, Canada  
May 2024 - Current

- **SMILE: Semantic Role Extraction**
  - Refined SPARQL-based ontology of social service impact models through developing advanced semantic entity extraction techniques, improving the match between social purpose organizations and client needs.
  - Developed corpus of over 800 organizations and their impact models, trained domain-specific language models (e.g., RoBERTa) to enhance the accuracy of extracting impact models from unstructured text.

### University of Toronto Scarborough

Data Analyst Intern

Toronto, Canada  
Jan 2024 - May 2024

- **CareerNavigator: LLM & Knowledge Graph-Based Job Recommender Engine**
  - Developed a Flask-based job recommender system for co-op students, enhancing job matching and employment opportunities.
  - Shifted from NLP text extraction to leveraging Gemini-Pro for prompt engineering, boosting keyword extraction accuracy by 60%.
  - Utilized Neo4j Aura to construct a job-related knowledge graph for job matching with graph search and node similarity algorithms.
- **Weekly Report Data Generation System Reconstruction Project**
  - Reverse-engineered 1000+ lines of Python scripts used for weekly report data generation to decode data manipulation processes.
  - Designed SQL-based data generation logic with UML and crafted 600+ queries for MySQL to the foundation of automated reporting.

## SKILLS

**Languages:** Python, SQL, Java, C, Bash, Cypher, SPARQL, Node.js

**Machine Learning Frameworks:** Pytorch, Tensorflow, Keras, spaCy, Scikit-learn, XGBoost, Matplotlib, OpenCV, CUDA

**LLM Frameworks:** Hugging Face, Langchain, Ollama

**Cloud Tools:** AWS, Google Cloud, FireBase, Neo4j AuraDB, MongoDB Atlas

**Development Tools:** Git, MySQL, GraphDB, Unix/Shell, Tableau, PowerBI, Markdown, Latex, Github, Jupyter Notebook, Anaconda