Yongan(Michael) Yu

yya040327@gmail.com | (438)-821-1823 | Linkedin | Github | Portfolio

EDUCATION

McGill University
August 2022 - April 2026
Montréal, QC

Bachelor in Computer Science Major & Geographical Information System Minor

Courses: Applied Machine Learning, Natural Language Processing, Artificial Intelligence, Database Systems, Algorithm Design, Software Design, Data Structure, Advanced Geographic Information Science

The University of Hong Kong
January 2025 - May 2025
Hong Kong, SAR

Exchange Student at the Department of Computer Science

PUBLICATIONS

- **WXImpactBench: A Disruptive Weather Impact Understanding Benchmark for Evaluating Large Language Models.**
  Yongan Yu, Qingchen Hu, Xianda Du, Jiayin Wang, Fengran Mo, & Renée Sieber
  Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL), 2025. (Accepted ACL findings)
- **From Recall to Reasoning: Automated Question Generation for Deeper Math Learning through Large Language Models.**
  Yongan Yu, Alexandra Krantz, & Nikki G. Lobczowski
  26th International Conference on Artificial Intelligence in Education (AIED), 2025. (Accepted - acceptance rate: 16%)
- **MaintainCoder: Maintainable Code Generation Under Dynamic Requirements.**
  Zhengren Wang*, Rui Ling*, Chufan Wang*, Yongan Yu*, Zhiyu Li, Feiyu Xiong, & Wentao Zhang.
  ArXiv. https://arxiv.org/abs/2503.24260 (Under review - NeurIPS 2025)
- **CodeFlowBench: A Multi-turn, Iterative Benchmark for Complex Code Generation**
  Sizhe Wang, Zhengren Wang, Dongsheng Ma, Yongan Yu, Rui Ling, Zhiyu Li, Feiyu Xiong, Wentao Zhang
  ArXiv. https://arxiv.org/abs/2504.21751 (Under review - NeurIPS 2025 D&B)
- **THiNK: Can Large Language Models Think-Aloud**
  Yongan Yu, Mengqian Wu, Yiran Lin, Nikki G. Lobczowski
  (Under review - EMNLP 2025)

*Equal Contributions

AWARDS & SCHOLARSHIPS

BMO Junior Responsible AI Scholars Award	$8,700 | Bank of Montreal (BMO) & McCAIS
Friends of the Libraries UG Award	$1,000 | Trenholme Dean of Libraries, McGill
Mobility Bursary for Exchanges 2025	$8,000 | Issued by Quebec government (PMICSE)
Research Assistantship Stipend 2024	$2,263 | Issued by Internal Social Sciences and Humanities Development (SSH)
OpenAI Researcher Access Program - Research Grants 2024	$5,000 | Issued by OpenAI Research Team
Undergraduate Student Research Award (USRA) 2023	$6,000 | Issued by Canadian Institutes of Health Research (CIHR)

RESEARCH EXPERIENCES

Research Intern
May 2024 - Present
Data Rescue Archives & Weather(DRAW), McGill Montréal, QC

- Under the supervision of Prof. **Renée Sieber**, created the first benchmark for evaluating LLMs' understanding of disruptive weather impacts, established a four-stage data construction pipeline and implemented multi-label classification and ranking-based QA tasks.
- As the 1st author, the paper was submitted to ACL 2025, and we proposed an open-source repository to help future society protect against vulnerabilities from disasters.

Research Intern
June 2024 - Present
Department of Education, McGill Montréal, QC

- Under the supervision of Prof. **Nikki Lobczowski**, led a two-phase research study investigating LLM-based question generation for mathematics education, examining how contextual information affects output quality via RAG, with a specific focus on creating questions across varying cognitive complexity levels using Bloom's Taxonomy and Webb's Depth of Knowledge framework.
- Building on these findings, currently collaborating with Mila to develop a broader benchmark for evaluating LLM's understanding and application in educational contexts.

Research Intern
Digital Data Design Institute (D³), Harvard
Januaray 2024 - April 2024
Boston, MA

- Under the supervision of Prof. **Jacqueline Lane**, published a case study on leveraging GenAI to enhance venture capital decision-making, exploring predictive analytics and real-time data integration for investment strategies.
- Developed an application that evaluated over 2,000 startup products using Tree-of-Thought reasoning and K-means clustering, achieving 79% accuracy in identifying circular economy ventures based on product value, material access, and processing levels.

Research Intern
Rosalind & Morris Goodman Cancer Institute
January 2023 - December 2023
Montréal, QC

- Under the supervision of Prof. **Luke McCaffery**, contributed to modeling cell simulation to investigate how macrophages destroy the membrane and affect cell mitosis in the epithelial layer.
- Compiled a parallel center mass graph window to track cell simulation and analyze them by the CPM (GGH) algorithm.

WORK EXPERIENCES

AI Research Fellow
Borealis AI, RBC
September 2024 - December 2024
Montréal, QC

- Processed and refined over **10GB** of raw OCR-scanned historical text data by implementing LDA for topic modeling and applying advanced techniques such as GPT-based filtering, Unicode normalization, and regex-based corrections to enhance data quality.
- Labeled and annotated data subsets from the historical weather dataset, running multiple open and closed-source models (e.g., GPT, Gemini, Claude, LLaMA, Mixtral, etc.) to perform binary classification on societal vulnerabilities and resilience strategies.

TECHNICAL SKILLS:

- **Programming:** Python, Java, C/C++, R, CSS, HTML, JavaScript, TypeScripy, SQL, Bash, XML
- **Tools**: LangChain, Scikit-learn, Git, Matplotlib, Pandas, NumPy, Selenium, Plotly, Spark, Linux
- **Frameworks**: TensorFlow, PyTorch, ReactJS, MongoDB, Django, NodeJS, FastAPI, AWS

SERVICE

Serves as ACL ClimateNLP 2025 Reviewer