ZACHARY YANG

RSTZZZ | zachary-y-647209103 | zachary.yang@mail.mcgill.ca

EDUCATION

Doctoral of Computer Science, McGill University — Expected 2025
Relevant Coursework: Natural Language Understanding with Deep Learning

Master of Computer Science, McGill University — 2020 - 2022
Relevant Coursework: Network Science, Distributed Systems, Natural Language Processing, Applied Machine Learning

Honours Bachelor of Computer Science, University of Toronto — 2015 - 2019
Co-operative Program in Software Engineering Stream | Graduated with High Distinction | Dean's List for all years

HIGHLIGHTED PUBLICATIONS

Towards Detecting Contextual Real-Time Toxicity for In-Game Chat, Z.Yang, N. Grenon-Godbou, R. Rabbany. In the proceedings of *Findings of the Association for Computational Linguistics: EMNLP 2023*

- Developed a real-time toxicity detection system that integrates NER and dialogue state tracking across multi-turn, multi-speaker conversation, thereby enhancing the content safety moderation for online chat systems.

Unveiling Identity Biases in Toxicity Detection : A Game-Focused Dataset and Reactivity Analysis Approach, J. Van Dorpe, Z.Yang, N. Grenon-Godbou, W. Grégoire. In the proceedings of *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP): Industry Track*

- Explores and mitigates inherent biases in NLP models to establish a more trustworthy, responsible and ethical content safety framework.

Online Partisan Polarization of COVID-19, Z.Yang, A. Imouza, K. Pelrine, S. Levy, J. Liu, G. Desrosiers-Brisebois, J. Godbout, A. Blais, R. Rabbany. In the proceedings of *2021 IEEE International Conference on Data Mining Workshops on Social Data Mining in the Post-pandemic Era (ICDMW-SDM) pp.893-901, IEEE 2021*

- Conducts comprehensive analysis of online partisan polarization, providing political scientists with critical insights to identify and examine divisive issues.

EXPERIENCE

NLP R&D Intern — May 2022 - August, 2023
Ubisoft La Forge — *Montreal, QC*

- Spearheaded the advancement of toxicity detection algorithms, resulting in a **significant improvement (+43%)** in the F1-score, and established **industry-leading player content safety systems**
- Pioneered an unsupervised learning project through integrated active learning and human-in-the-loop methodologies, advancing the **trust and safety strategy important to chat moderators**.
- Research on detecting and preventing toxicity within in-game chat using language models, with **two papers in EMNLP 2023** and a **presentation in Ethical Gaming 2024**.

Graduate Research — Jan 2021 - Present
Complex Data Lab — *Montreal, QC*

- Developed **scalable** classifiers and tools using machine learning and data mining techniques to measure partisan polarization for **large-scale data (over 80K users and 30M posts)**
- **Collaborated with cross-domain research teams** to correlate this measure with existing COVID-19 epidemiology data and political events to investigate the potential causes and impacts of polarization
- Designed the first text-based measurement of partisan polarization on social media in the context of COVID-19 across time and between states in the United States and Canada, resulting in **one paper in ICDMW**, one presentation at McGill and **one paper in IEEE VIS**.

SKILLS

Programming Languages	Python 3, C#, Java
ML Packages	HuggingFace, PyTorch, Numpy, Pandas, NLTK, Sklearn, Bokeh, Scipy, SpaCy
DB	Oracle, PostgreSQL, MongoDB, Firebase
Version Control	Azure, TFS, GitHub, GitLab
Scripting	PowerShell, Batch, Linux
Soft Skills	Time Management, Problem-solving, Attention to Detail, Adaptability

PUBLICATIONS UNDER REVIEW

Web Retrieval Agents for Evidence-based Misinformation Detection, J. Tian, H. Yu, Y. Orlovskiy, M. Rivera, Z.Yang, J. Godbout, K. Pelrine. Submitted to *COLM, 2024*

- Leveraging LLMs, web-search, retrieval-agumented generation, statement decomposition to fact check statements.

OTHER PUBLICATIONS

Party Prediction for Twitter, K. Pelrine, A. Imouza, Z.Yang, G. Desrosiers-Brisebois, S. Levy, J. Tian, , C Amadoro, A. Blais, J. Godbout, R. Rabbany. In the proceedings of *International AAAI Conference on Web and Social Media, 2024*

Game On, Hate Off: A Study of Toxicity in Online Multiplayer Environments, Z.Yang, N. Grenon-Godbou, R. Rabbany. Presentation in *Ethical Games Conference, 2024*

When does Continuous Learning for BERT make sense?, Z.Yang. In the proceedings of *Proceedings of the Canadian Conference on Artificial Intelligence, 2023*

Open, Closed, or Small Language Models for Text Classification, Z.Yang, Y. Hao, K. Pelrine, J. Godbout, R. Rabbany. Preprint published on *ArXiv, 2023*

COVID-19 Partisan Polarization and Toxicity, Z.Yang, K. Pelrine, A. Imouza, G. Desrosiers-Brisebois, S. Levy, J. Tian, J. Godbout, R. Rabbany. Poster presented at *McGill School of Computer Science 50th Anniversary, 2022*

Activity Based Party Prediction for Twitter, K. Pelrine*, A. Imouza* , G. Desrosiers-Brisebois*, S. Levy*, J. Tian*, Z. Yang*, A. Feizi*, A. Blais, JF. Godbout , R. Rabbany. In the *American Political Science Association Meeting (APSA), 2022*

OPPVIS: Visualizing Online Partisan Polarization of COVID-19, Z.Yang, A. Imouza, K. Pelrine, S. Levy, J. Liu, G. Desrosiers-Brisebois, J. Godbout, A. Blais, R. Rabbany. In the proceedings of *2021 IEEE Visualization & Visual Analytics (VIS 2021), IEEE, 2021*

Ebbs and Flows of Polarization During a Political Campaign, K. Pelrine, A. Imouza, G. Desrosiers-Brisebois, Z.Yang, S. Levy, A. Feizi, J. Liu, A. Blais, J. Godbout, R. Rabbany. In the *American Political Science Association Meeting (APSA), 2021*

OTHER EXPERIENCE

DevOps Engineer (Ministry of Education, *Toronto, ON*) Jan 2018 - Sep 2020

- Implemented CI/CD pipelines for builds and releases, significantly enhancing code quality feedback and **reducing deployment errors by over 90%**, demonstrating the ability to improve and maintain high-quality software systems.
- Designed and implemented a new automation Framework in C#, that resulted in a **cost savings of $200k in yearly licensing fees**, reflecting the capacity to develop cost-effective solutions.

IT QA (Ministry of Education, *Toronto, ON*) Sep 2016 - Dec 2017

- Orchestrated the implementation and scheduling of automated testing using HP UFT and Selenium, helping in cutting page loading times by 60% and **significantly boosting user experience**
- Upgraded the automation framework with HP UFT and Selenium to incorporate over 40 new interactions, including web element status verification and report validations, ensuring comprehensive testing and robust automation