

Karolina Tchilinguirova

k.tchiling@gmail.com | karolina.codes | linkedin.com/in/karolina-tchilinguirova | github.com/KarolinaTchiling

Education

Bachelor of Science (Honours), Computer Science <i>York University – Lassonde School of Engineering, Toronto, ON</i>	Expected: Aug 2026
Bachelor of Science (Honours), Specialization in Geological Sciences <i>Queen's University – Geological Sciences and Geological Engineering, Kingston, ON</i>	Completed: Dec 2019

- Awards: Three-time recipient of the Lassonde Undergraduate Research Award (\$24,300)
- GPA: 3.32/4.0

- Certificate in Geographic Information Sciences (GIS)

Experience

Data Analytics Co-op, Hydro One – Reliability & Analytics	May 2025 – Current
<ul style="list-style-type: none">• Cleaned, validated, and verified the integrity of large-scale operational outage datasets used for enterprise-wide internal and external reliability reporting.• Designed and implemented a Python-based data validation and automation pipeline leveraging graph theory, statistical analysis, and exploratory data analysis (NetworkX, Pandas, NumPy) to replace a legacy Microsoft Access workflow, enabling scalable processing and automating 70%+ of the data cleaning process, resulting in multiple days per month of time savings.• Reconstructed Hydro One's electrical network as a Python graph-based model, enabling automated topology-aware validation of outage customer counts that previously required manual device tracing and customer reconciliation.• Leveraged this network reconstruction to build Hydro One's first visual circuit connectivity analysis tool, adopted across multiple teams to validate distribution and transmission network relationships.• Developed Power BI dashboards and optimized SQL-based analytics workflows to support operational decision-making, ad-hoc analysis, and modernization of legacy reliability reports.• Applied software engineering best practices, including modular architecture, version control, and documentation, using Git and GitLab to ensure maintainable, production-quality analytics code.	
Data Science Researcher, DARE! Lab at York University	May 2024 – April 2025
<ul style="list-style-type: none">• Conducted large-scale citation analytics on 96,000+ software engineering papers to quantify gender-based citation bias using statistical modeling and time-series analysis (2009–2024) to evaluate temporal trends in citation behavior.• Designed data ingestion, cleaning, and validation pipelines, integrating metadata from third-party sources including Web of Science, CrossRef, OpenAlex, Semantic Scholar, and Gender-API.• Applied statistical analysis to compare observed vs. expected citation behavior and used bootstrap resampling to assess the statistical significance of citation bias patterns across multiple data subsets.• Implemented reproducible analytics workflows using Python, R, MongoDB, and Pandas, and communicated findings through published preprints and conference presentations.	
Geo-spatial Machine Learning Researcher, GNSS Lab at York University	May 2023 – April 2024
<ul style="list-style-type: none">• Initiated a research project investigating the feasibility of using supervised machine learning techniques to improve smartphone-based navigation accuracy using large-scale, real-world GNSS datasets.• Conducted preliminary model exploration using k-nearest neighbors (KNN), Random Forest classifiers, and Support Vector Machines implemented in Scikit-Learn to classify high- and low-quality positioning measurements based on spatial correlation patterns.• Extended the analysis by training and cross-validating Random Forest models using GNSS signal-based features, including carrier-to-noise density ratio (C/N_0), elevation angle, and pseudorange consistency, for LOS/NLOS signal classification.• Collaborated on the development and evaluation of machine-learning-assisted measurement covariance scaling methods within a Precise Point Positioning (PPP) framework, resulting in a peer-reviewed conference publication demonstrating significant reductions in smartphone positioning error in multipath-affected environments.	

Publications and Preprints

Improving Smartphone Positioning by Adapting Measurement Noise Covariance using Machine Learning
Peer-reviewed Conference Paper, ION GNSS+ 2024 • DOI:10.33012/2024.19761

Biases in Gendered Citation Practices: An Exploratory Study and Some Reflections on the Matthew and Matilda Effects
Preprint, 2024 • ARXIV:2410.02801

Gendered Citation Practices in Software Engineering Reference Lists: A Replication Study
Preprint, 2025 • SSRN:5200880

CiteFairly: Toward the Improvement of Gendered Citation Practices in the Software Engineering Field
Preprint, 2025 • SSRN:5450972

Technologies

Languages: Python, Java, C++, C#, R, SQL, JavaScript, TypeScript **Data Science & ML:** Pandas, NumPy, Scikit-Learn, NetworkX, Matplotlib, Power BI **Databases & Platforms:** PostgreSQL, PostGIS, SQL Server, MongoDB, Firebase, Azure, GCP
Frameworks & Tools: Node.js, Next.js, Three.js, React, Express, Spring Boot, Flask, Linux, Git, GitHub, GitLab, JUCE, Unity