Chapin S. Korosec, PhD

chapinSKorosec@gmail.com | chapinkorosec.ca | LinkedIn

Summary

Applied Data Scientist and AI/ML Engineer specializing in healthcare and large-scale structured data systems. Experienced in building end-to-end data pipelines, machine learning models, and simulation frameworks to support decision-making in clinical and operational environments. Strong background in data wrangling, statistical modeling, and AI systems (including transformer-based and agentic workflows), with a focus on translating complex datasets into deployable, user-facing solutions. Experienced in communicating technical findings to non-technical stakeholders through reports, visualizations, and presentations in technical, clinical and policy-facing environments. Experienced in developing simulation and scenario-based models to support healthcare system capacity planning and policy evaluation, and working with real-world healthcare data to support system-level planning and decision-making. Focused on building usable models that inform real-world healthcare decisions.

Technical Skills

- **Languages:** Python, R, SQL, C++, Matlab
- **Data Engineering:** Data ingestion, ETL pipelines, data cleaning, transformation, feature engineering **ML/AI:** Scikit-learn, PyTorch, Hugging Face Transformers, Random Forests, Generative Models (cVAE, GANS), classification and regression models
- **AI Systems:** Retrieval-Augmented Generation (RAG), embedding-based retrieval (FAISS), agentic workflows, LLM integration, automation pipelines, PEFT
- **Data Tools:** Pandas, NumPy, DuckDB, data.table
- **Methods:** Simulation modeling (microsimulation, agent-based, discrete-event, compartmental), statistical modeling, epidemiological modeling (SEIR), nowcasting, forecasting, predictive analytics, model validation and sensitivity analysis, synthetic data generation and validation
- **Visualization:** Shiny, Plotly, custom dashboard development
- **Other:** Git, API integration, reproducible workflows, data de-identification

Communication & Professional Skills

- Technical report writing and documentation for scientific, clinical, and policy-facing audiences.
- Delivered 30+ conference presentations, invited talks, and academic lectures across interdisciplinary domains.
- Experience collaborating with clinicians and healthcare researchers on real-world data analysis and modeling projects.
- Communicated complex modeling results to both technical and non-technical stakeholders, including public health and clinical audiences.
- Produced data visualizations, dashboards, reports, and summaries to support decision-making and knowledge translation.

Professional Experience

Technical Lead, EpiSense, Toronto **July 2025 – Present**

- Building a real-time disease surveillance and analytics platform integrating user-reported and public health data streams.

- Designed and implemented end-to-end data pipelines (ingestion, cleaning, transformation, storage) for structured and semi-structured health data.
- Developed machine learning and statistical models to estimate disease burden and forecast trends across geographic regions.
- Built automated analytics workflows and decision-support dashboards for real-time monitoring and reporting.
- Integrated AI-driven components including retrieval-based systems and automated data interpretation pipelines.
- Developed scenario-based analytics to support assessment of disease trends and potential healthcare system pressures.

Program Manager, Modeling and Economics Research Network (ModERN) Nov 2023 – July 2025
Canadian Immunization Research Network (CIRN)

- The focus of (ModERN) is to conduct epidemiological analyses, mathematical modeling, and economic analyses to study the costeffectiveness and population-level effectiveness of public health interventions.
- Contribute to management of seminars and annual budget and writings of grants and research proposals.

Senior Machine Learning Research Scientist, York University April 2021 - Nov 2023
Mathematics and Statistics Department

- My research focused on immunogenicity of COVID-19 infection and vaccination and predicting SARS-CoV-2 mutation dynamics.
- Awarded University of Toronto 'Artificial Intelligence for Public Health' fellowship.
- Authored successful $30,000 COVID-19 Immunity Task Force grant.
- First Canadian to be Awarded the 2024 Michelson Postdoctoral Prized Lectureship in its 27 year history.
- Developed nonlinear mixed-effects models to characterize longitudinal immune response kinetics (IFNg, IL2, IgG, IgA, neutralization) following mRNA SARS-CoV-2 vaccination using multi-timepoint clinical cohort data (n=139) from long-term care populations.
- Quantified heterogeneous immune trajectories (growth vs decay regimes) via parameterized doubling times and half-lives, enabling predictive linkage between early T-cell responses and long-term neutralization capacity.
- Integrated multi-modal immunological datasets including PBMC-derived ELISpot cytokine responses, serological antibody titers (ELISA), and live-virus neutralization assays to quantify cellular–humoral coupling and decay dynamics.
- Modeled the impact of pre-existing cross-reactive immunity on vaccine response, incorporating demographic and exposure heterogeneity in high-risk populations (long-term care staff in high-density regions).

Academic Work Experience

Adjunct Professor, University of Guelph July 2025 – Present

- Designed and delivered upper-year courses in linear and nonlinear optimization (Operations Research), receiving strong student evaluations.
- Built and deployed **PharmaTrace**, a web-based ML-enabled analytics platform for Canadian pharmaceutical import/export data (1988–2025).

- Conducted applied research on infectious disease dynamics, including mathematical modeling of measles spread in Ontario and Alberta using real-world public health data.
- Invited to present national research contributions under the INSPIRE grant at the WE-SPARK Health Institute conference

Selected Data Science & AI Projects

- Developed AI-driven pipelines combining structured data with language models for automated insight generation and reporting.
- Built retrieval-augmented generation (RAG) systems for querying biomedical and clinical datasets using embeddings and semantic search.
- Designed forecasting and simulation models to support decision-making under uncertainty in epidemiological systems.
- Implemented automated data workflows for integrating multiple heterogeneous data sources into unified analytical frameworks.

Education

Ph.D., Physics — Simon Fraser University
B.Sc. (Honours), Physics — McMaster University

Selected Achievements

- 2024 Michelson Postdoctoral Prized Lectureship (first Canadian recipient)
- NSERC Postdoctoral Fellowship
- 22+ **peer-reviewed publications**, including a *Nature Communications* paper and 3 cover articles with significant media coverage.