

# NORMAN BUI

+1 613-983-3217 | [normanbui@gmail.com](mailto:normanbui@gmail.com) | [linkedin.com/in/normanbui23](https://www.linkedin.com/in/normanbui23) | [github.com/nbui23](https://github.com/nbui23)

## EDUCATION

- Carleton University** | *Bachelor of Computer Science (Honours), 3.9/4.0 GPA* Expected April 2026
- Awards: Deans' Honours List, President's Scholars Scholarship, Chalmers Jack Mackenzie Scholarship
  - Relevant Coursework: Data Structures & Algorithms, Discrete Math, Web Apps, Operating Systems, Object-Oriented Software Engineering, Statistics, Programming Paradigms, Human-Computer Interaction (UI/UX), Databases

## EXPERIENCE

- Public Health Agency of Canada** | *Junior Data Scientist - Python, ETL, Web Scraping* Sep 2024 - Present
- Building ETL pipelines for web scraping using Python.

- Bedarra Corporation** | *Student Researcher - Python, LLM, RAG* Aug 2024 - Present
- Researched LLM and RAG systems to develop AI assistants for teaching niche programming languages (Factor, K).
  - Developed a RAG system for Factor programming language queries, increasing query response accuracy by 40%, through implementation of FLAN-T5 language model and FAISS vector store.

- Kongsberg Geospatial** | *Software Engineer Intern - C++, Qt, Python, GIS* May 2024 - Aug 2024
- Built GIS-focused demos with C++, Qt, Git, and TerraLens SDK for ATC and defense applications.
  - Implemented advanced map rendering (LiDAR, WMTS, S-57), object tracking, performance visualization (CPU, GPU, RAM), and UI enhancements for TerraLens demo, significantly boosting client satisfaction.
  - Developed a proof of concept script to handle ATC software's log processing with Python, **reducing processing time from days to seconds**.
  - Aggregated various VFR and terrain data sources using Python scripting, **reducing processing time by 40%**.

- Health Canada** | *Software Developer Intern - Python, ML, ETL, Web Scraping* Sep 2023 - Dec 2023
- Developed Python solutions for chemical assessment, leading to improved risk data aggregation and analysis.
  - Built random forest classifiers with pandas, NumPy, SciPy, and scikit-learn to accurately predict chemical toxicity.
  - Restored web scraping scripts for chemical data aggregation using BeautifulSoup4 and lxml, **increasing risk data search functionality by 53.33%**.
  - Refactored federated search system and added session management, automatic retries, and dynamic rate limiting, significantly enhancing system maintainability and reliability.
  - Developed automated workflow for extracting semantic triples from scientific articles using CoreNLP and Stanza, streamlining the collection and organization of qualitative data into a Neo4J knowledge graph.

- Department of National Defence** | *Junior Software Developer - Python, LLM, RAG* Jun 2023 - Nov 2023
- Built a LLM RAG proof of concept with Flask and GPT-3.5 to simulate conversation with a local document database.
  - Implemented a NLP pipeline with tokenization, context analysis, and named entity recognition using NLTK.
  - Used SQLite3 and FAISS for efficient storage and embedding retrieval, enabling fast response times.

- NAV Canada** | *Software Engineer Intern - C++, Qt, PowerShell, GIS* May 2023 - Aug 2023
- Developed C++ solutions for various ATC applications, resulting in improved performance and reliability.
  - Implemented optimized data management strategies using C++ and Qt to achieve a **60% reduction in directory clutter** and **improve automated analysis solutions by 30%** in internal data extraction tool.
  - Optimized code and resolved various bugs in ATC software, **improving overall graphical performance by 6%**.
  - Automated network configurations using PowerShell, **reducing setup time from 30 minutes to seconds**.

## PROJECTS

- SocialSaplings** | *HTML/SCSS, JavaScript, Bootstrap, Node, Express, Firebase (Firestore, Auth)* [Devpost](#)
- Led the development of a reforestation web platform, securing **1st place overall** at KurioHacks: March Edition.

## SKILLS

**Languages:** C/C++, Python, Java, SQL, JavaScript, PHP

**Frameworks & Databases:** Flask, Express, Bootstrap, WET, MySQL, SQLite3, PostgreSQL, Firebase, MongoDB, Neo4J

**Libraries:** pandas, NumPy, matplotlib, scikit-learn, unittests, spaCy, OpenAI, Sentence Transformers, BS4, lxml

**Development Tools:** Git, Node, Ubuntu, Qt Creator, Postman, Balsamiq, Agile Methodology, SDLC, OOP