NORMAN BUI

+1 613-983-3217 | normanbui@cmail.carleton.ca | linkedin.com/in/normanbui23 | github.com/nbui23

EDUCATION

Carleton University | *Bachelor of Computer Science (Honours),* ***3.9/4.0 (A) CGPA*** Expected April 2026

- Awards: Deans' Honour List, President's Scholars, Chalmers Jack Mackenzie, and Henry Marshall Tory Scholarships
- Relevant Coursework: Data Structures & Algorithms, Discrete Math, Web Apps, Operating Systems, Object-Oriented Software Engineering, Statistics, Programming Paradigms, Human-Computer Interaction (UI/UX), Databases, QA

EXPERIENCE

Junior Data Scientist | *Public Health Agency of Canada - Python, LLM, MDM* Sep 2024 - Present

- Collaborated with WHO, Imperial College London, and Health Canada to research, implement, and optimize LLMs for enhancing literature reviews in epidemiological research, as part of a Master Data Management system

Software Engineer Intern | *Kongsberg Geospatial - 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) and object tracking in 2D and 3D, performance visualizations for CPU, GPU, RAM, and various UI enhancements, **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%**

Data Scientist Intern | *Health Canada - Python, ML, NLP, 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

Junior Software Engineer | *Department of National Defence - Python, LLM, RAG* Jun 2023 - Nov 2023

- Built a 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

Software Engineer Intern | *NAV Canada - 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**

Junior Web Developer | *Correctional Service Canada - C#, .NET, Web App, QA* May 2022 - Apr 2023

- Improved victim services web application through front-end accessbility testing and bug fixes, increasing user satisfaction for **over 100 daily users**

RESEARCH & PROJECTS

RAG Systems for Obscure Languages | *Python, RAG, LangChain, FAISS, Sentence Transformers* GitHub

- Researched and developed RAG systems to answer obscure programming language queries @ Bedarra Corporation

SocialSaplings | *HTML/SCSS, JavaScript, Bootstrap, Node, Express, Firebase (Firestore, Auth)* Devpost

- Led the development of a reforestation web platform, securing **1st place overall** at KuriusHacks: March Edition
- Implemented a tool using various APIs to recommend optimal tree species based on a user's geodata
- Developed a tree planting visualization tool with Google Maps API to showcase global reforestation impact
- Added profiles, posts, likes, and comments, using Firebase for data management and authentication

SKILLS

Languages: HTML/CSS, C/C++, C#, Python, Java, SQL, JavaScript, PHP | English, French, Vietnamese
Frameworks & Databases: Flask, Express, MySQL, SQLite3, PostgreSQL, Firebase, MongoDB, Neo4J, JUnit, Selenium
Libraries: pandas, NumPy, matplotlib, scikit-learn, unittests, spaCy, Sentence Transformers, BS4, lxml, HuggingFace
Development Tools: Git, Node, Ubuntu, Qt Creator, Postman, Balsamiq, Agile Methodology, SDLC, OOP, TDD, Cucumber