BRUNO FONKENG

Ph.D. Researcher | Research Software | C/C++ Systems | ML Infrastructure

bruno@poisedre.com | bruno.fonkeng@und.edu | GitHub | LinkedIn | Google Scholar

ACADEMIC PROFILE

Computer Science Ph.D. researcher at the University of North Dakota with experience in **synthetic cybersecurity data, probabilistic generative modeling, trustworthy machine learning, reproducible scientific software, high-performance computing, and C/C++ systems development**. Research and engineering interests connect scientific rigor with dependable infrastructure, performance-aware implementation, and a long-term specialization in compiler and runtime engineering.

EDUCATION

Ph.D. in Computer Science

Aug 2025 – Present

University of North Dakota, Grand Forks, ND

Research focus: synthetic cybersecurity data, probabilistic and generative models, malware classification, trustworthy evaluation, reproducibility, and high-performance computing.

M.S. in Computer Science

Aug 2023 – May 2025

University of North Dakota, Grand Forks, ND

Graduate study in machine learning, artificial intelligence, cybersecurity, data engineering, secure applications, APIs, predictive modeling, and high-performance computing.

B.S. in Physics

Oct 2017 – Aug 2020

The University of Bamenda, Cameroon

Foundation in mathematical reasoning, scientific modeling, laboratory analysis, statistics, experimentation, and first-principles problem solving.

RESEARCH EXPERIENCE

Ph.D. Researcher

Aug 2025 – Present

University of North Dakota, Computer Science / ICCC Laboratory

Conduct doctoral research on synthetic cybersecurity data and the conditions under which probabilistic and generative models improve downstream malware classification.

- Develop GenCyberSynth, a research framework for training, sampling, comparing, and evaluating cybersecurity-data generators.
- Study GAN, VAE, diffusion, autoregressive, masked autoregressive, restricted Boltzmann, Gaussian-mixture, and related probabilistic approaches.
- Evaluate generated data using KID, MS-SSIM, balanced accuracy, macro F1, macro AUPRC, diversity measures, and downstream utility.
- Build reproducible HPC workflows using fixed seeds, experiment manifests, SLURM, artifact tracking, automated aggregation, figures, and publication-ready reports.
- Investigate balanced and class-selective synthetic-data policies under controlled real-data and generation budgets.

Research Assistant – TruNorth

Completed 2025

University research project supported by John Deere

Contributed to an applied decision-support initiative involving ROI calculation, requirements analysis, validation, system design, and technical documentation.

- Translated operational requirements into measurable inputs, calculations, outputs, and user-facing workflows.
- Supported validation, software-design decisions, applied research, and interdisciplinary documentation.

TEACHING EXPERIENCE

Graduate Teaching Assistant

Multiple academic terms

University of North Dakota

Supported undergraduate instruction, assessment, consultation, and technical feedback in systems programming, computer architecture, cybersecurity, computer networking, and information assurance.

- Assisted students with C programming, pointers, memory, file I/O, processes, debugging, systems interfaces, and command-line development.
- Reviewed technical submissions and provided structured feedback on correctness, software design, security, evidence, and technical communication.
- Supported laboratory, project, grading, and consultation activities across multiple computing courses.

COURSES SUPPORTED AS TEACHING ASSISTANT

CSCI 330 – Systems Programming — *Multiple terms*
CSCI 370 – Computer Architecture — *Teaching support*
CSCI 389 – Computer and Network Security — *Spring 2025*
CSCI 327 – Data Communications — *Fall 2025*
CSCI 290 – Cybersecurity and Information Assurance — *Teaching support*

PUBLICATIONS AND MANUSCRIPTS

First-author peer-reviewed conference paper — *ISICN 2025, accepted*
Peer-reviewed research contribution in intelligent computing and networking. Full citation should match the final proceedings and Google Scholar record.

When Does Synthetic Data Help Malware Classification? — *Active conference and journal manuscript pipeline*
Examines when synthetic malware data improves downstream classification under controlled model families, generation budgets, and evaluation protocols.

Selective Synthetic Data Policies for Malware Classification — *Active conference and journal manuscript pipeline*
Studies balanced and class-selective augmentation strategies using class-aware evaluation, controlled data budgets, and downstream classification outcomes.

PRESENTATIONS AND POSTERS

IEEE CARS Conference Presentation — *2024, presented*
Delivered a technical research presentation to an academic and professional audience.

Red River Valley ACS Research Poster — *2024, presented*
Communicated experimental methods and findings through a research poster and technical discussion.

SELECTED TECHNICAL PROJECTS

GenCyberSynth — *Python, PyTorch, SLURM, HPC, generative models*
Reproducible framework for synthetic malware-data generation, model comparison, downstream evaluation, artifact tracking, aggregation, and publication reporting.

C Systems Mastery — *C, GCC, Make, GDB, Valgrind, sanitizers, Linux*
Structured systems-programming roadmap covering safe input, arrays, pointers, modular interfaces, testing, memory, debugging, data structures, and operating-system concepts.

MetricForge — *C++, Python, CMake, Python bindings, ML metrics*
Open-source machine-learning metrics project with a C++ core, Python bindings, tests, examples, clean public interfaces, and education-focused documentation.

C CLI Lab — *C, Linux, POSIX, command-line interfaces, file I/O*
Extensible suite of Unix-inspired tools built to deepen understanding of streams, files, processes, arguments, text processing, directories, and POSIX behavior.

VotingSphere — *Django, Flutter, secure application design*
Secure online voting platform developed as an applied software project.

Fraud Detection API — *Python, FastAPI, scikit-learn*
Machine-learning API for fraud detection using a Random Forest model and service-oriented deployment.

ADDITIONAL PROFESSIONAL EXPERIENCE

Project Dynamo Contributor
2026 – Project paused
Handshake AI
Contributed to AI-training and evaluation tasks involving precise specifications, Python automation, Docker, structured JSON outputs, and automated verification.

- Developed and refined technical tasks involving log processing, file-system operations, structured results, and test-driven verification.
- Evaluated instruction precision, reproducibility, output schemas, failure conditions, and agent behavior.

TECHNICAL SKILLS

Programming Languages: C; C++; Python; SQL; Bash / shell scripting

Systems Engineering: Memory and pointers; file I/O; processes; modular interfaces; data structures; concurrency foundations; Linux; POSIX; computer architecture

Machine Learning and Research: PyTorch; generative models; synthetic data; malware classification; experiment design; ML metrics; reproducibility; data pipelines

Tools and Infrastructure: Git; GitHub; GCC; Make; CMake; GDB; Valgrind; sanitizers; Docker; SLURM; HPC; WSL; Jupyter; VS Code

Compiler and Runtime Foundations: Lexing and parsing concepts; abstract syntax trees; intermediate representations; virtual machines; execution models; optimization concepts

Software and Data: REST APIs; FastAPI; JSON; data processing; testing; automation; documentation

SELECTED GRADUATE COURSEWORK

- CSCI 532 – High Performance Computing and Paradigms
- CSCI 543 – Machine Learning
- CSCI 551 – Security for Cloud Computing
- CSCI 557 – Computer Forensics
- CSCI 587 – Ethical Hacking
- CSCI 589 – Application Layer Security
- DATA 527 – Predictive Modeling
- DATA 530 – Artificial Intelligence
- DATA 532 – Applied Machine Learning
- DATA 540 – Data Visualization
- MATH 442 – Linear Algebra
- CSCI 999 – Dissertation Research
- EECS 500 – Graduate Seminar

RESEARCH INTERESTS

- Synthetic cybersecurity data
- Probabilistic generative modeling
- Trustworthy machine learning
- Malware classification
- Reproducible scientific computing
- High-performance computing
- ML systems and AI infrastructure
- Compiler and runtime systems
- Performance engineering

PROFESSIONAL PROFILES

Portfolio: poisedre.com | GitHub: edgemindstudio | LinkedIn | Google Scholar