William Nixon

williamnixon20@gmail.com williamnixon20

Research Interests

ML-for-systems, systems-for-ML, operating systems, storage systems, distributed systems, and databases.

Education

Computer Science, Bandung Institute of Technology **Jul 2021 - Exp. Jul 2025**
#1 Engineering University in Indonesia. GPA: 3.95/4.00. *Bandung, Indonesia*
TOEFL: 115 [30 R, 29 L, 27 S, 29 W]

Publication

W. Nixon*, R. A. O. Sinurat*, A. I. Kistijantoro, H. S. Gunawi. **"Data Drift Mitigation In Learned Caches"**. *Manuscript available upon request. In Preparation.*

H. Lovenia, R. Mahendra, S. M. Akbar, L. J. Miranda, J. Santoso, ... **W. Nixon**, ... S. Cahyawijaya. **"SEACrowd: A Multilingual Multimodal Data Hub and Benchmark Suite for Southeast Asian Languages"**. *In Proceedings of the Empirical Methods in Natural Language Processing (**EMNLP), 2024.***

Research Experience

Research on Workload Drift in Caching Systems **Nov 2023 – Present**
Undergraduate Researcher *Remote*

- Collaborated with Prof. Haryadi S. Gunawi (University of Chicago) and Prof. Achmad Imam Kistijantoro (Bandung Institute of Technology) to improve **miss ratio** in caching systems by addressing **workload drift** in learned caches.
- Investigated how workload drift affects **machine learning-based caching eviction models**, aiming to improve cache performance under changing workload patterns.
- Created a processing pipeline to **analyze large-scale block I/O traces** from over 800 nameservers provided by Tencent and Alibaba, automating the processing of terabytes of data to characterize and identify patterns of drift.
- Implemented various strategies to enhance caching performance by applying **drift mitigation algorithms** to a learned cache **GL-Cache**, **reducing miss ratio by up to 5%**

Open Source Research Experience (OSRE) **Jun 2024 – Aug 2024**
Undergraduate Researcher *Remote*

- A competitive **NSF-funded summer fellowship** contributing to reproducibility research held by the UC Santa Cruz.
- Worked with Dr. Sandeep Madireddy (Argonne National Lab) to reproduce and benchmark **drift mitigation algorithms** in **machine learning-based I/O admission model**, which decides whether to admit or failover I/O requests by modeling an SSD's state as busy or normal.
- Reproduced and implemented algorithms such as **Matchmaker**, **DriftSurf**, and **Accuracy Updated Ensemble (AUE)** from their respective research papers, despite the lack of available source code.
- Integrated mentioned algorithms into an I/O admission model, which successfully **improved P99 latency** than baseline under drifting conditions
- Packaged the research pipeline into a **Chameleon Trovi artifact** to ensure reproducibility and facilitate future research efforts.

Research on Continual Learning in Systems **Jan 2023 – Nov 2023**
Undergraduate Researcher *Remote*

- Collaborated with Prof. Haryadi S. Gunawi (University of Chicago) on applying **continual learning** techniques to large-scale system performance modeling.
- Addressed the challenge of **data drift** in machine learning models by applying continual learning methods to datasets from **Argonne National Laboratory's HPC system trace** and the **Google Borg Scheduler**.
- Integrated and benchmarked **7+ continual learning algorithms**, under changing data distributions, achieving up to a **30% improvement in accuracy** compared to naive retraining methods.

SeaCrowd Project **Nov 2023 – Feb 2024**
Undergraduate Researcher *Bandung, Indonesia*

- Participated in a joint collaboration to collect NLP datasets for **Southeast Asian languages**, which are underrepresented in the research community due to data scarcity.
- Implemented standardized dataloaders for seamless integration with **HuggingFace's platform**, facilitating public dataset use by researchers and developers.
- Curated, catalogued, and standardized over **30+ datasets** spanning various language modalities, from text to speech.

Work Experience

Bandung Institute of Technology **Jan 2023 – Jun 2024**
Teaching Assistant *Bandung, Indonesia*

- **Probability and Statistics (Spring 2024)**: Head TA. Led a team in designing and grading homework, class project for **150+ students**.
- **Discrete Mathematics (Fall 2023)**: Designed and graded quizzes and homework assignments for over **150+ students**.
- **Introduction to Python Programming (Spring 2023)**: Proctored and mentored class labs consisting of **20+ students**.

TEDx ITB **Jan 2023 – Apr 2023**
Backend Developer *Bandung, Indonesia*

- Served as the sole backend developer for TEDx ITB, an annual event hosting over **200 participants**.
- Developed the **ticketing system** and implemented backend logic to ensure smooth event operations, contributing to raising **20 million IDR** in revenue.
- **Tech Stack**: Next.js (Node.js & React), Strapi CMS, SendGrid.

Bukit Vista Hospitality Services **Mar 2022 – Dec 2022**
Backend Developer *Remote*

- Developed a **backend calendar system** for a mobile app to allow property partners to view their properties' occupancy in a Google-calendar-like manner.
- Automated various businesses processes in **finance and HR**, reducing time otherwise consumed by up to **90%**.
- **Tech Stack**: Node.js, Python, AWS, Integromat.

Highlighted Projects

Simple OS | ***C, QEMU*** |

- Developed a basic **operating system kernel** from scratch, with support for interrupts, keyboard input, and FAT32 file system management.
- Implemented key filesystem operations such as `ls`, `cd`, `mkdir`, `whereis`, `cat`, `rm`, `cp`, and `mv`, providing full CRUD functionality.

3D WebGL Editor | ***JavaScript, React, WebGL*** |

- Built a **3D model editor** using WebGL primitives (w/o library) for rendering hollow and articulated models with real-time animation transformations such as translation, rotation, and scaling.
- Implemented interactive features including model save/load functionality, camera control with multiple projection types, and shading with material properties like diffuse, specular, and texture mapping.
- Supported advanced features like dual canvas views, GPU-based picking, tweening, and post-processing effects.

TCP over UDP | ***Python*** |

- Implemented **TCP protocol over UDP** (unreliable socket), under unreliable network conditions (packet delay, corruption, duplication, and loss).
- Developed core features including **three-way handshake**, **file transfer**, error-correcting codes (**Hamming**), and proper connection termination.
- Built a **Tic-Tac-Toe game** on top that can run across multiple devices in the same network.

Technical Skills

Programming Languages: C, C++, C#, CSS, HTML, Java, JavaScript, Python, SQL
Operating Systems: Windows, WSL, Linux, MacOS
Machine Learning: Jupyter Notebook, PyTorch, TensorFlow, Scikit-learn, Pandas
Databases and Systems: SQLite, PostgreSQL, MongoDB, Cassandra, DuckDB
Systems: Apache Spark, RabbitMQ, Redis, FEMU
Cloud Computing: Google Cloud, AWS, Chameleon Cloud Testbed, Chameleon Trovi
Web Development: React.js, Node.js, Nest.js, Next.js , Django, Vercel
Mobile Development: Android, React Native
DevOps: GitHub CI/CD, Vercel, Docker
Misc: LaTeX, Make

References

Haryadi S. Gunawi
Associate Professor
University of Chicago
haryadi@cs.uchicago.edu

Achmad Imam Kistijantoro
Associate Professor
Bandung Tech
imam@itb.ac.id

Ayu Purwarianti
Associate Professor
Bandung Tech
ayu@itb.ac.id

Rinaldi Munir
Associate Professor
Bandung Tech
rinaldi@itb.ac.id