

Minguk Choi

mgchoi@dankook.ac.kr | <https://min-guk.github.io/>

RESEARCH INTERESTS

- **Systems for ML:** Resource-efficient Training/Serving on Cloud/Edge, AI Agents
- **ML for Systems:** Learned Index, Learned Cache, Learned Scheduler

EDUCATION

Dankook University

- Master of Science in AI-based Convergence

Yongin, Korea
Mar 2023 – Aug 2024

Dankook University

- Bachelor of Science in Software Science
- GPA: 4.1 / 4.5 (**3rd Place** in Department)

Yongin, Korea
Mar 2017 – Feb 2023

PUBLICATIONS

International Conference

[SIGMOD 2024] Can Learned Indexes be Built Efficiently? A Deep Dive into Sampling Trade-Offs

Top Conference

Minguk Choi, Seehwan Yoo, and Jongmoo Choi

[\[Paper\]](#) [\[Slides\]](#) [\[Poster\]](#) [\[Code\]](#)

International Journal

[Electronics 2023] An Empirical Study of Segmented Linear Regression Search in LevelDB

Ramadhan Agung Rahmat, Minguk Choi, Yoojin Chung, and Jongmoo Choi.

Selected Domestic Conference (+7 more)

[KCC 2024] Analysis of RMI Using CPU-Optimized Search Algorithms

Best Paper Award

Yejin Oh, Minguk Choi, Boseung Kim, Yongjie Zhu, Seehwan Yoo, and Jongmoo Choi

[KCC 2024] Breakdown Internal Operations in Updatable Learned Index

Best Presentation Award

Suhwan Shin, Minguk Choi, Nakyeong Kim, Seehwan Yoo, and Jongmoo Choi

ON-GOING PROJECTS

Federated Learning Plan under Privacy Constraints: Compile the optimal federated runtime plan for end-to-end ML pipelines (e.g., data preparation, debugging, and training) using a cost model based on the different privacy constraints of data, workers, and tasks in Apache SystemDS

Exploring the Design Space for SIMD Acceleration in Learned Indexes: Introduce novel approaches that accelerate learned indexes by leveraging SIMD and data parallelism in internal operations (e.g., error-bound estimation, model-biased insert). Additionally, it extends the SIMD design space of index structures from horizontal to vertical vectorization.

EXPERIENCES

Remote Research Internship

DAMS Lab (Data Management for Data Science Laboratory)

Technische Universität Berlin, Germany

Aug 2024 – Present

- Adviser: [Prof. Matthias Boehm](#)

Conducted remote research from Korea based on Apache SystemDS, focusing on compiler and runtime backend for local, distributed, and federated environments. Currently working on a federated learning project under privacy constraints.

Student Researcher

System Software Laboratory

Dankook University, Korea

Jun 2021 – Aug 2024

- Adviser: [Prof. Seehwan Yoo](#) and [Prof. Jongmoo Choi](#)

Researched Linux Kernel (CPU Scheduler, Memory Allocator, Block I/O Stack), Key-Value Store (Skiplist, KV Cache, Compaction, SST File, Bloom Filter), and Learned & Traditional Index Structures (Sample Learning, Semi-Ordered Structure, SIMD)

Compulsory Military Service
Air Force Operations Command Service & Support Group
Honorably discharged as a sergeant after completing mandatory service

Korean Air Force
Jan 2019 – Oct 2020

HONORS AND AWARDS

Best Research Award , Dankook University Top 3 Research Achievements Among PhD and Master's Students in Science at Dankook University	Aug 2024
Best Paper Award , Korea Computer Congress 2024	Jun 2024
Best Presentation Award , Korea Computer Congress 2024	Jun 2024
Certificate of Appreciation , Korea Computer Congress 2024	Jun 2024
Academic Excellence Award , Dankook University Graduated with high honors (3rd in Software Science department)	Feb 2023
Battalion Lead Soldier , Korean Air Force Represented and supervised soldiers, and supported recruit adaptation	Mar 2020

PROFESSIONAL SERVICES

Availability & Reproducibility Committee, ACM SIGMOD 2024
Evaluated 2 SIGMOD 2024 artifacts: one for Availability and another for both Availability & Reproducibility

OPEN-SOURCE CONTRIBUTION

Apache SystemDS , An end-to-end ML system with hybrid local and distributed execution compilation Conducting a Federated Learning project under privacy constraints; implemented the roll function in local, distributed (Spark), and federated versions; fixed incorrect varID reference when requesting federated data; corrected a bug that miscounted element count when writing mtx headers	Contributer
BASIL , Benchmark for Sampling Applied Learned Indexes Created a benchmark that fairly evaluates the build time and lookup performance of 8 state-of-the-art sampling applied learned and traditional indexes on 12 sorted datasets (SIGMOD 2024 Artifact)	Owner
LevelDB WIKI , Analysis document about LevelDB (Key-Value Store) Co-authored a document with 13 undergraduate students analyzing LevelDB, detailing theory, code, and including results from benchmark experiments and the YCSB tuning contest.	Owner
YCSB-CPP , Yahoo! Cloud Serving Benchmark (YCSB) written in C++ Fixed underflow and compilation bugs, and added support for LevelDB property options	Contributer
Uftrace , Function graph tracer for C/C++/Rust/Python Introduction to How to Analyze Key-Value Stores (LevelDB, RocksDB) Using uftrace	Wiki Contributor

TEACHING ASSISTANT

Operating Systems , Dankook University Developed and graded assignments on 8 different CPU scheduler simulations, concurrent data structures (Queue, BST), and Ext2 file system forensics. Coding assignments were auto-graded using Google Test and MOSS. Answered student questions and proctored exams.	Spring 2024
Index Structure , Dankook University Delivered a 4-hour introductory lecture on traditional and learned index structures. Mentored 5 papers for domestic conferences, resulting in one best paper award and one best presentation award.	Winter 2023
System Programming , Dankook University Created and graded assignments on file management and shell development (myLs, myCreate, myCat, myCopy, myShell), answered student questions, and proctored exams.	Fall 2023
Operating System Practice , Dankook University Delivered a 40-hour lecture on operating system. Developed and supervised exercises on basic tools (vim, ssh, gcc, gdb), file handling (myCat, myCreate), concurrency problems (Dining Philosophers, Readers-Writers), and shell implementation. Answered student questions, created, proctored, and graded exams.	Summer 2023
Key-value Store , Dankook University Delivered a 6-hour introductory lecture on Key-value Stores (LevelDB), mentored 4 papers for domestic conferences, organized a YCSB tuning contest, and co-authored open-source documentation with students.	Summer 2022

LANGUAGE AND TECHNICAL SKILLS

Languages: Korean - Native, English - TOEFL 104/120 (RC: 29, LC: 29, SP: 22, WR: 24)

Programming Languages: C/C++, Java, Python

PERSONAL INTERESTS

Fitness: Running, Weight lifting

Cooking: Korean Cuisine, Japanese Cuisine, Grilling

REFERENCES

[Prof. Jongmoo Choi](#) choijm@dankook.ac.kr

Software Science, Dankook University, Korea

[Prof. Seehwan Yoo](#) seehwan.yoo@dankook.ac.kr

Mobile Systems Engineering, Dankook University, Korea

[Prof. Seong-je Cho](#) sjcho@dankook.ac.kr

Software Science, Dankook University, Korea

[Prof. Matthias Boehm](#) matthias.boehm@tu-berlin.de

Big Data Engineering, Technische Universität Berlin, Germany