

Minguk Choi

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

RESEARCH INTERESTS

Scalable and Efficient Systems for Machine Learning:




- Training:** Cost-based optimization of federated learning execution plans
- Retrieval:** Hybrid vector-relational databases and indexing structures
- Inference:** Optimization of large language model inference with KV caching

EDUCATION

University of Wisconsin–Madison PhD in Computer Science	Madison, WI, USA Starting Sep 2025
Dankook University MS in AI-based Convergence	Yongin, Korea Mar 2023 – Aug 2024
Dankook University BS in Software Science (3rd place in Department)	Yongin, Korea Mar 2017 – Feb 2023

PUBLICATIONS

International Conferences & Journals

- I2. (SIGMOD '24) Can Learned Indexes Be Built Efficiently? A Deep Dive into Sampling Trade-offs **Best Artifact Award**
[Minguk Choi](#), [Seehwan Yoo](#), and [Jongmoo Choi](#)    [\[Paper\]](#) [\[Slides\]](#) [\[Poster\]](#) [\[Code\]](#)
- I1. (Electronics '23) An Empirical Study of Segmented Linear Regression Search in LevelDB
[Ramadhan Agung Rahmat](#), [Minguk Choi](#), [Yoojin Chung](#), and [Jongmoo Choi](#) [\[Paper\]](#)

Domestic Conferences (+7 more)

- D2. (KCC '24) Analysis of RMI Using CPU-Optimized Search Algorithms **Best Paper Award**
[Yeojin Oh](#), [Minguk Choi](#), [Boseung Kim](#), [Yongjie Zhu](#), [Seehwan Yoo](#), and [Jongmoo Choi](#) [\[Paper\]](#)
- D1. (KCC '24) Breakdown Internal Operations in Updatable Learned Index **Best Presentation Award**
[Suhwan Shin](#), [Minguk Choi](#), [Nakyeong Kim](#), [Seehwan Yoo](#), and [Jongmoo Choi](#) [\[Paper\]](#)

Technical Reports

- T2. (ARI @ SIGMOD '24) Reproducibility Report for ACM SIGMOD 2024 Paper: "Optimizing Distributed Protocols with Query Rewrites"
[Anwesha Saha](#), [Minguk Choi](#), [Qiaolin Yu](#), and [David CY Chu](#) [\[Paper\]](#)
- T1. (ARI @ SIGMOD '24) Reproducibility Report for ACM SIGMOD 2024 Paper: "Can Learned Indexes be Built Efficiently? A Deep Dive into Sampling Trade-offs"
[Yihao Liu](#), [Yuanhui Luo](#), and [Minguk Choi](#) [\[Paper\]](#)

HONORS AND AWARDS

Best Artifact Award, ACM SIGMOD 2024 Recognized for providing fully reproducible, flexible, and portable research artifacts	Jun 2025
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 Awarded for a collaborative research with my mentee, Yeojin Oh (B.S. student)	Jun 2024
Best Presentation Award, Korea Computer Congress 2024 Awarded for a collaborative research with my mentee, Suhwan Shin (M.S. student)	Jun 2024
Academic Excellence Award, Dankook University Graduated with high honors (3rd place in software science department)	Feb 2023

RESEARCH EXPERIENCES

Research Collaboration DIAS Lab (Data-Intensive Applications and Systems Laboratory) – Collaborator: Dr. Kyoungmin Kim An algorithm to optimize multi-vector top-k queries combined with complex relational filters in vector-relational hybrid databases.	EPFL, Switzerland Mar 2025 – Present
Research Internship DAMS Lab (Data Management for Data Science Laboratory) – Advisor: Prof. Matthias Boehm Optimized federated execution plans for end-to-end ML pipelines in Apache SystemDS under privacy constraints, using a dynamic programming-based cost model.	TU Berlin, Germany Aug 2024 – Present
Research Assitant System Software Lab – Advisor: 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)	Dankook University, Korea Jun 2021 – Jul 2024

PROFESSIONAL SERVICES

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

TALKS

Can Learned Indexes be Built Efficiently? A Deep Dive into Sampling Trade-Offs Top Conference Session @ Korea Computer Congress 2024, Jeju, Korea, Jul, 2024 ACM SIGMOD/PODS Conference 2024, Santiago, Chile, Jun, 2024 ICAN (ICT Challenge & Advanced Network of HRD) Workshop @ Dankook University, Yongin, Korea, Dec, 2023

OPEN-SOURCE CONTRIBUTION

Apache SystemDS , An end-to-end ML system with hybrid local and distributed execution compilation Contributed to federated learning under privacy constraints in Apache SystemDS; implemented vector roll functions across execution modes (local, Spark, federated) and resolved bugs related to data alignment and aggregation.	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	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	Contributer

PROFESSIONAL EXPERIENCES

Index Data Structure Study Group Leader & Research Mentor Delivered a 4-hour lecture on index structures and mentored 5 student projects, resulting in one best paper and one best presentation award at domestic conferences.	Dankook University Jan 2024 – Jun 2024
Key-value Store Study Group Leader & Research Mentor 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.	Dankook University Jul 2022 – Dec 2022

TEACHING EXPERIENCES

Operating System Practice

Teaching Assistant

Dankook University
Summer 2023, Winter 2024, Spring 2025

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.

Operating Systems

Teaching Assistant

Dankook University
Spring 2024

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.

System Programming

Teaching Assistant

Dankook University
Fall 2023

Created and graded assignments on file management and shell development (myLs, myCreate, myCat, myCopy, myShell), answered student questions, and proctored exams.

OTHER EXPERIENCES

Compulsory Military Service

Air Force Operations Command Service & Support Group

Korean Air Force
Jan 2019 – Oct 2020

Honorably discharged as a sergeant after completing mandatory service

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

REFERENCES

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

Dept. Software Science, Dankook University, Korea

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

Dept. Mobile Systems Engineering, Dankook University, Korea

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

Dept. Big Data Engineering, Technische Universität Berlin, Germany