

DONGWON KIM

✉ dkim010@gmail.com | 📱 dkim010 | 💻 dongwon-kim-68235a35 | 🎓 Google Scholar

Summary

ML platform engineering leader with 10+ years of experience building large-scale data and production ML systems. Currently serving as Director of ML Engineering (Principal IC track), driving company-wide technical strategy for ML platforms, reliable model serving, and emerging LLM/Agent workloads. Proven track record of scaling engineering organizations and delivering high-performance infrastructure in fintech and consumer domains.

Work Experience

ML/AI Platform in Viva Republica

Financial super app, Toss, has 15M+ MAU in Sep. 2023.

Seoul, Korea, May 2022 - Present

- Director of ML Engineering (Nov. 2025 - Present)

- ML Platform Team Leader (Oct. 2024 - Present)

- Interim AI Platform Team Leader (Aug. 2025 - Dec. 2025)

• Leadership & Platform Ownership

- **Scaled ML Organization:** Expanded leadership from a single 5-person team to coordinating 20+ engineers across multiple teams, unifying architectural standards and platform strategy.
- **Strategic Vision:** Spearheaded the technical evolution of a company-wide ML platform supporting recommendation, ads, and generative AI/Agent workloads.
- **Operational Excellence:** Consolidated fragmented initiatives into a unified platform, significantly reducing duplicated efforts and boosting engineering velocity.

• ML/AI Platform

- **Scalable Inference Engine & Lifecycle Management:** Architected a large-scale inference platform handling 10K+ RPS and hundreds of model inferences per request, while centralizing the control plane for hundreds of models, reducing deployment lead time by 70% (to approx. 30% of previous levels).
- **Feature Store & Standardized Framework:** Developed an in-house Feature Store guaranteeing P95 latency under 15ms and RecoKit ensuring training-serving consistency, accelerating model time-to-market from days to hours.
- **Vector Retrieval:** Built and operated dual-datacenter vector database clusters with CPU/GPU hybrid indexing, enabling real-time ingestion and low-latency search over tens of millions of embeddings for semantic search and targeting.
- **GPU Efficiency:** Built and optimized a Kubernetes GPU platform by reclaiming idle GPU capacity and applying Volcano with selective gang scheduling, significantly reducing resource allocation failures and supporting 100+ engineers.
- **Automated DAG Builder:** Engineered an Airflow-based factory that enabled 100+ non-data engineers to operate 500+ production DAGs daily, featuring automated SLA enforcement, lineage tracking, and infrastructure rate-limiting.

Finance Data Platform of NAVER FINANCIAL

Financial services has 13M+ MAU in Mar. 2021.

Sungnam, Korea, Mar. 2020 - May 2022

- Technical Leader (Jul. 2020 - May 2022)

• Financial Data/ML Platform

- **Secure Data Infra:** Designed and operated a Hadoop-based secure data platform supporting 200+ internal users, enabling reproducible analytics in regulated financial environments.
- **ETL Framework:** Built a config-driven data ingestion to standardize data collection, validation, and versioning across heterogeneous sources.
- **Causal Inference:** Established an experiment analysis framework to evaluate model-driven campaigns, including causal impact analysis and feedback loops for iterative model improvement

FLO Music Service of Dreamus Company

Dreamus Company, a subsidiary of SK Telecom, has 2M+ memberships in Jan. 2020.

Seoul, Korea, Dec. 2018 - Mar. 2020

- Recommendation Team Leader (Jan. 2019 - Mar. 2020)

• Music Recommender System

- **Cloud Transformation:** Built a cloud-native big data platform from scratch, replacing legacy on-premise systems and supporting 100+ engineers.
- **Personalized Playlist:** Implemented embedding-based recommendation pipelines that increased playlist adoption rates from under 5% to 27% within six months.
- **MIR Integration:** Built a continuous mood feature extraction system using Music Information Retrieval (MIR) techniques, integrating deep learning-based audio signals into recommendation models.

Music Streaming Service of NAVER

Sungnam, Korea, Aug. 2015 - Nov. 2018

NAVER's music streaming service has 0.8M+ memberships in 2018.

• Music Recommender System

- **Personalized Station:** Developed a real-time music station recommender incorporating continuous user feedback, achieving a 20% user adoption rate through high-precision similarity search.
- **Discovery Engine:** Implemented a song discovery system using hybrid collaborative and content-based filtering, leading to a 10% increase in long-tail song discovery.

Education

M.Sc. in Computer Science

Mar. 2011 - Aug. 2015

YONSEI UNIVERSITY,

Seoul, Korea

"Content-centric energy management for mobile devices" (Advisor: Prof. Hojung Cha, Mobile Embedded System Lab.)

B.S. in Computer Science

Mar. 2007 - Feb. 2011

YONSEI UNIVERSITY

Seoul, Korea

Selected Publications

Deep Learning (3 papers)

- B. Jeon, A. Kim, C. Kim, **D. Kim**, J. Park, J. Ha, "Music Emotion Recognition via End-to-End Multimodal Neural Networks," *Proceeding of the 11th ACM Conference on Recommender Systems (RecSys 2017)*, Aug. 2017

System Engineering (12 papers)

- **D. Kim**, N. Jung, Y. Chon, H. Cha, "Content-Centric Energy Management of Mobile Displays," *IEEE Transactions on Mobile Computing (TMC)*, DOI: 10.1109/TMC.2015.2467393
- C. Yoon, **D. Kim**, W. Jung, C. Kang, H. Cha, "AppScope: Application Energy Metering Framework for Android Smartphone using Kernel Activity Monitoring," *Proceeding of 2012 USENIX Annual Technical Conference (USENIX ATC)*, Jun., 2012.