

Myeong Heon (Sean) Choi

📍 Seoul, South Korea ✉ mooner92@kakao.com ☎ +82 10-6668-1871 🌐 mooner92.info
in myeong heon (sean) choi 🏠 mooner92

Summary

Site Reliability / Systems Engineer focused on distributed systems, Kubernetes, and large-scale infrastructure reliability. Published cluster-scheduling research, bare-metal-to-production cluster operations, and two years coaching algorithms and competitive programming.

Skills

Languages: C++, C, Python, Bash, SQL, TypeScript
Infrastructure: Kubernetes, Docker, Linux, Nginx / Ingress, Prometheus, Grafana, Loki, Cloudflare (Tunnel / Zero Trust)
Cloud: Google Cloud (GCP), AWS, Oracle Cloud

Experience

AI Researcher (Contract), Infrastructure & Platform Sejong, South Korea
Korea Environment Institute (KEI) Apr 2026 – Present

- Operate a multi-server GPU compute fleet (dual NVIDIA A40 nodes) serving LLM inference and per-user JupyterLab for institute researchers
- Migrated a production server from Ubuntu 18.04 to 24.04 with zero data loss, validating RAID6 and running an approximately 155 GB rsync backup to NAS beforehand
- Built the observability stack (Prometheus + Grafana + Loki), containerized internal services with Docker, and added WireGuard VPN and Cloudflare Zero-Trust access

Co-Founder Manchester, UK
Branchify Sep 2024 – Nov 2025

- Co-founded a secondhand-marketplace web app; led market research (TAM/SAM/SOM), Figma prototyping, and the front-end / back-end split

SW Quality Certification Researcher Bundang, South Korea
Telecommunications Technology Association (TTA) Jul 2024 – Dec 2024

- Authored the AI-model bias verification checklist for AI medical devices, operationalizing ISO/IEC TR 24027 into checkable items on dataset composition, class balance, and representativeness
- Authored the security control checklist by mapping ISO/IEC 27001 / 27002 / 27032 (and AAMI TIR57, IEC 62304) into device-level controls across communication security, authentication/authorization, input validation, and key management

Instructor, Algorithms Daejeon, South Korea
JR Coding Academy Jun 2022 – Jun 2024

- Taught data structures & algorithms in C++ and Python; coached the KOI competitive-programming team and produced preliminary- and final-round qualifiers each year

Software Engineer Intern Daejeon, South Korea
Engstok Jun 2023 – Aug 2023

- Built an e-commerce estimation chatbot (NestJS) and migrated the backend from Ubuntu to the company's Windows server with customer-DB integration

Projects

TPU-Aware Pod Scheduling on Kubernetes github.com/mooner92/tpuserv



- Built a bare-metal Kubernetes cluster (four Raspberry Pi nodes with Coral Edge TPU accelerators) and used node-affinity scheduling to route accelerator tasks to TPU-equipped nodes, cutting task execution time approximately 70% vs. traffic-based default scheduling

- Instrumented the cluster with Prometheus and Grafana and managed certificates with Cert-Manager
- Tools: Kubernetes, containerd, Nginx Ingress, TFLite

KEIAdminSuperv — On-Prem RAG for Internal Regulations

github.com/mooner92/KEIAdminSuperv [🔗](#)

- Built and operate an on-premises RAG service answering staff questions over 363 internal regulation documents (approximately 4,830 embedded chunks), with multi-turn chat, per-answer source citations, and isolated dev / prod environments — fully on-prem on dual GPUs behind Cloudflare Zero Trust
- Hardened the citation guarantee (every answer cites its source article or refuses) from 80.6% to 100%, and improved strict retrieval accuracy (Hit@1) from 60.0% to 82.9% with a cross-encoder reranker, measured on a golden evaluation set
- Tools: FastAPI, Chroma, KURE-v1 embeddings, Ollama (Qwen3.5-9B), Next.js, SQLite

Publications

Scheduling Techniques for Improving the Efficiency of TPU Work in a Cluster Environment

Published Jun 2024

Myeong Heon Choi, Hyun Suk Oh, Ji Hoon Park

[Korean Institute of Information Scientists and Engineers \(KIISE\)](#) [🔗](#)

Research on Improving Pothole Object Detection Accuracy Through Similar-Object Data

Published Jun 2024

Hyun Suk Oh, Myeongheon Choi, Ji Hoon Park

[Korean Institute of Information Scientists and Engineers \(KIISE\)](#) [🔗](#)

Proposal of a Korean History Education Application Based on Kubernetes Clustering

Published Nov 2023

Seul Lee, Bonyeong Lee, Myeongheon Choi

[I410-ECN-151-24-02-089052651](#) [🔗](#)

A Study on Chatbot for a Safe Harbor

Published Nov 2023

Youngmin Kang, Sangwook Kim, Hyunsuk Oh, Myeongheon Choi

[I410-ECN-151-24-02-089053890](#) [🔗](#)

Education

Konyang University

Mar 2019 – Feb 2025

B.S. in Computer Engineering (Medical IT Engineering)

- GPA: 3.94/4.5
- **Coursework:** Operating Systems, Computer Architecture, Data Structures & Algorithms, Computational Theory, Databases, Linux Systems

DSC University

Mar 2023 – Feb 2025

B.S. in Mobility Software / AI Convergence

- GPA: 4.04/4.5
- **Coursework:** Machine Learning Engineering, Operating Systems, Computer Vision, Embedded Software

Certifications

- **Engineer Information Processing** — Acquired Sep 2024
- **Engineer Information Security** — Acquired Dec 2024
- **SQL Developer (SQLD)** — Acquired Jun 2024