



Shan Hu

☎ (+86) 195-2295-1401 • ✉ hs@kmoon.fun • 🌐 blog.kmoon.fun
🗣 @Kmoons_ • 📷 @kmoonn




🎓 Education

- | | |
|---|--|
|  Wuhan University of Technology
Software Engineering GPA: 4.0/5.0 Rank: 20%
First-class Academic Scholarship; 1 pending Invention Patent; 2 Software Copyrights | Recommended Master
Sep. 2024 – Jul. 2027 |
|  Wuhan University of Technology
Information Engineering GPA: 3.9/5.0 Rank: 20% CET-6
Outstanding Graduate, Excellent Graduation Thesis; University-level Outstanding Student Leader, College-level Merit Student | Bachelor
Sep. 2020 – July. 2024 |


🏆 Awards

- | | |
|--|--|
| • National Innovation Training Program | Project Leader Jul. 2025 |
| • 12th China Software Cup | National Second Prize Aug. 2023 |
| • 14th Service Outsourcing Innovation Competition | Central China Third Prize Jun. 2023 |
| • 13th National E-commerce Triple Innovation Challenge | Hubei First Prize Apr. 2023 |


💼 Internships

 **ByteDance** **Shanghai**
Test Development Intern Global Monetization Product and Technology May. 2026 – Present
Oversaw TikTok GMPT third-party creative QA covering advertisers, partners & creators; focused on SDT and AI Testing toolchain development.

- **Creative BQA Agent**: AI backend testing agent
- **AICO-Test Plugin**: AI Copilot testing plugin

 **KuaiShou** **Beijing**
Test Development Intern Commercial Technology Department Feb. 2026 – Apr. 2026
Supported Kuailiao and CRM QA & test dev for outbound and risk control teams; built efficiency analysis, alert diagnosis and auto-inspection capabilities.

- **Performance Analysis Platform Development**: Designed and implemented a multi-dimensional QA workflow for demand risk, scheduling, team and personal efficiency. Built full-link data processing with Function Calling + Pydantic for stable JSON Schema output. Adopted strategy registry pattern for high extensibility and maintainability.
- **Complaint Diagnosis Platform Development**: Developed backend alert services and iterated platform features. Implemented 3-level intent classification for automated complaint handling. Built rate-limiting counters with Redis to avoid duplicate/missing alerts and enabled real-time robot push for fault detection.
- **UI Automation for Chat App**: Established UI automation inspection based on KRunner, covering core business flows. Integrated with cloud testing platform for daily scheduled regression, identified and fixed 4 critical interaction defects, improving release stability.
- **Bug2Case Association Skill**: Independently built bidirectional bug-test case linkage using multi-dimensional matching rules. Connected data sources to support auto recommendation and batch binding. Optimized the defect regression closed-loop and reduced missed testing risks.

 **iFLYTEK** **Hefei**
Test Development Intern EdTech Platform Department Oct. 2025 – Dec. 2025
Conducted test development for EdTech products including Spark grading device and Zhixuewang APP, focusing on automation tooling, environment monitoring and stability optimization.

- **Answer Sheet Auto-rendering Tool**: Developed with PyPDF & ReportLab; leveraged FastAPI + Spark LLM to auto-generate exam content with simulated handwriting. Cut 8h manual preparation for a class down to within 5 mins.
- **Production Data Sync Tool**: Built pipeline to import real grading data into test env; one-click data replication boosted data collection efficiency by 70%+ and end-to-end verification by 90%.
- **Disk Auto-clean Utility**: Implemented disk monitoring & log auto-cleanup with real-time alert via IFlytek Webhook, improving environment self-recovery and branch deployment success rate.
- **API Automation**: Designed 200+ JMeter regression cases covering 100+ core APIs across 5 modules; integrated into pre-release regression to enhance test coverage and validation speed.

BPEG Intelligent Retrieval and Q&A Platform

<https://ragent.kmoon.fun>

Backend Development Java 17 + Spring Boot 3 + PostgreSQL + Milvus + RocketMQ + React 18 Jun. 2025 – Sep. 2025

Intro: Built enterprise knowledge Q&A platform, responsible for multi-source retrieval, model routing, full-link tracing and high-concurrency governance.

- **MultiChannel Parallel Retrieval Engine:** Constructed SearchChannel+PostProcessor two-layer architecture to fix single-recall missing problem; used CompletableFuture + dedicated thread pool to run keyword & vector retrieval in parallel, followed by deduplication and rerank postprocessing. Boosted overall recall by 30%+, Top-K relevance +25%, complex query accuracy from 62% to 87%.
- **Fault-tolerant Model Routing Layer:** Developed unified gateway for Chat/Embedding/Rerank models, equipped with priority routing and three-state circuit breaker. Realized automatic model switching upon failure with below 500ms switch latency, effectively avoid core Q&A service breakdown caused by single model exception.
- **Custom Asynchronous Full-link Tracing:** Designed TraceRun/TraceNode tracking model and @RagTraceNode annotation to propagate trace context across thread pool and MQ. Realized end-to-end monitoring for streaming response, async data sync and RocketMQ consumption, greatly lowering troubleshooting cost of abnormal slow requests.
- **Distributed Fair Queue & Rate Limiting:** Implemented Redis+Lua based fair queue system for long-lived SSE requests, supporting task cancel, timeout recycle and resource release. Controlled traffic surge and model resource leakage, keeping queue timeout ratio under 5%.

WHUT Competition Integrated Management System

<https://contest.kmoon.fun>

Full-stack Development Java 11 + Spring Boot 2 + MySQL + Redis + MyBatis + JWT + Vue 3 Dec. 2024 – Apr. 2025

Intro: Full-featured campus competition management platform for event arrangement, award statistics and workload settlement, stably serving over 40k teachers and students for more than one year.

- **CAS-based SSO Authentication:** Integrated school CAS center via ticket verification to implement cross-platform single sign-on; adopted Sa-Token distributed Token to deliver stateless session management and unified campus identity interoperability.
- **Dynamic Multi-level Data Permission:** Leveraged MyBatis interceptor to dynamically inject filter SQL, designed role-department-custom three-layer permission framework with five configurable data ranges, preventing illegal cross-department data access with good extensibility.
- **High-efficiency Stream Excel Parsing:** Implemented event-driven SAX import framework to avoid OOM on large file upload; realized batch persistent storage and exception capture, improving bulk Excel import efficiency by 200%-300%.

Skills

Java: Proficient in Java and collection frameworks; familiar with JVM, GC, concurrent programming, thread pools and deadlock analysis.

Python: Skilled in Python for API development and automation scripting; experienced with FastAPI, Flask and LangChain.

Data Structures & Algorithms: Master common data structures and core algorithms: two pointers, backtracking, dynamic programming and sorting.

Computer Networks: Familiar with TCP/IP model and protocols like HTTP/HTTPS, TCP, UDP, IP and DNS.

Operating System: Understand OS fundamentals, process and thread management; proficient in Linux commands and Shell scripts.

Database: Experienced with MySQL and Redis; grasp indexes, transactions, locks and typical cache problems.

Development Frameworks: Proficient in Spring ecosystem and MyBatis; familiar with IoC, DI, AOP and Bean lifecycle.

AI: Hands-on experience with LLM, Agent, RAG, MCP and prompt engineering; skilled in Claude Code and Trae.

Design Patterns: Familiar with Singleton, Factory, Proxy, Observer, Producer-Consumer and other design patterns.

Others: Proficient in IDEA, Git and technical documentation writing.