

Zhang Jia

18292060445 · plus_zzz@163.com · 2002-09
Homepage: pluszzz.top
GitHub: github.com/Pluszzz



Education

2024 - Present	Northwest University 211 & Double First-Class	Software Engineering / M.S.
	• "Huawei Cup" Graduate Math Modeling — National Second Prize	
2020 - 2024	Xi'an University of Posts & Telecommunications	Information & Computing Science / B.S.
	• National Math Modeling Contest — Provincial Second Prize • CET-4	

Projects

AgentAI	github.com/Pluszzz/AgentAI	2025-09 ~ 2026-02
Overview: Enterprise RAG intelligent agent platform covering the full pipeline from document ingestion to intelligent Q&A — intent recognition, question rewriting, multi-channel retrieval, session memory, model routing with fault tolerance.		
Stack: SpringBoot + MyBatis Plus + Milvus + Redis + Redisson + Apache Tika + Sa-Token		
• Multi-channel retrieval engine (intent-directed + global vector + post-processing pipeline)		
• LLM-based question rewriting and decomposition for automatic context completion		
• Session memory compression (sliding window + auto-summary); Model routing with three-state circuit breaker		
• Tree-structured intent recognition with confidence-based clarification; MCP tool integration framework		
• Redis ZSET + Lua sliding window rate limiter; TransmittableThreadLocal for async context propagation		
Short Link Platform	github.com/Pluszzz/ShortLinkPlatform	2025-05 ~ 2025-08
Overview: SaaS short link system providing efficient, secure URL management with analytics for enterprise and individual users.		
Stack: Spring Boot + Spring Cloud Alibaba + MySQL + Redis + RocketMQ + ShardingSphere + Sentinel		
• MQ idempotency via Redis; Cache penetration protection with double-check locking and null-value caching		
• Bloom filter optimization; Sentinel QPS-level rate limiting with graceful degradation		
• RocketMQ async consumption for peak-shaving; "Update DB then delete cache" strategy		
• Redisson distributed read-write locks; ShardingSphere shard routing for targeted queries		

Skills

AI Coding

AI Agent workflows. Skilled with Codex, Claude Code for code generation, debugging, and automated development.

Java & JVM

Java concurrency (JUC, AQS, CAS), thread pool tuning. JVM memory model, GC (G1/CMS), OOM diagnosis.

Backend Frameworks

Spring Boot / Spring Cloud Alibaba. Nacos, OpenFeign, Sentinel for service discovery and traffic governance.

Databases

MySQL index design, execution plans, MVCC, transaction isolation. Slow query optimization.

Caching & Middleware

Redis data structures, persistence. Cache solutions. Distributed locks, rate limiting, Bloom filters.

LLM Applications (RAG)

Enterprise RAG pipeline: Milvus, multi-channel recall, question rewriting, session memory, model routing.

Engineering

Git, Maven, modular development.