

KESHAV BUDHIRAJA

SOFTWARE ENGINEER | FULL-STACK, AI SYSTEMS & MOBILE ARCHITECTURE

✉ keshavbudhiraja4321@gmail.com | Gurugram, Haryana, India | 🌐 keshavbudhiraja.com | 🔗 LinkedIn | 🐙 GitHub

PROFESSIONAL SUMMARY

Software Engineer specializing in full-stack web platforms, autonomous AI systems, and native mobile applications. Experienced in architecting multi-agent code review state machines (LangGraph, Python AST), enterprise Hybrid RAG engines (pgvector, BM25, RRF, Cross-Encoders), and interactive document verification studios (Next.js, React, TypeScript), alongside shipping 2 commercial mobile titles on Google Play with 99.85% crash-free stability. Certified Oracle Cloud Architect Associate focused on low-latency systems and production scalability.

TECHNICAL SKILLS

Core Languages: TypeScript, Python, Java, Kotlin, JavaScript, SQL, C#, C++.

AI Systems & Retrieval: LangGraph, Multi-Agent Systems, Hybrid RAG (pgvector + BM25 + RRF), Cross-Encoder Re-Ranking, Python AST, Pydantic v2, Ragas, Multi-LLM Gateways (Gemini, OpenAI, Claude, Groq).

Full-Stack & Backend: Next.js, React, FastAPI, Node.js, Express.js, PostgreSQL (pgvector), Supabase, Redis, REST APIs, Tailwind CSS.

Mobile & Game Engineering: Android SDK, Kotlin Coroutines, StateFlow, MVVM, Room DB, Custom 2D Canvas, Unity Engine, ProGuard/R8.

Cloud, DevOps & Security: AWS (S3, EC2, Lambda), Cloudflare (Pages, R2), GCP, Azure Blob, Oracle Cloud (OCI Certified), Docker, CI/CD (GitHub Actions), Linux CLI.

WORK EXPERIENCE

Puzzler Games | Software Engineer (Founding Engineer)

[puzzlergames.online](#)

- Production Lifecycle & Architecture:** Engineered and launched 2 commercial mobile applications on Google Play from zero to production, orchestrating modular MVVM architectures, release signing, and automated Gradle build pipelines.
- Concurrency & High Reliability:** Architected asynchronous state management using Kotlin Coroutines and StateFlow; decoupled intensive compute from the UI thread to eliminate frame drops and sustain a **99.85% crash-free session rate**.
- Offline Persistence & Edge Web Infrastructure:** Designed an indexed SQLite/Room persistence layer (**<12ms** read latency) and deployed the studio web hub on Cloudflare Pages with TLS 1.3 security and automated Git CI/CD hooks.

ENGINEERING PROJECTS

AetherDoc Studio — Document Intelligence & HITL Platform ([GitHub](#))

Next.js, React, TypeScript, Supabase, PDF.js, AWS S3

- Dual-Engine Extraction:** Built a Universal Multi-LLM BYOK Gateway with automated JSON recovery paired with a 100% offline coordinate-sorted PDF.js and native DOCX XML extraction engine.
- 60 FPS Verification Studio:** Engineered an interactive Human-in-the-Loop canvas using a reactive state store for proportional bounding-box calibration, live box drawing/deletion, inline overrides, and an immutable audit trail.
- Multi-Cloud & Database Sync:** Implemented zero-dependency AWS Signature V4 request signing for direct Cloudflare R2, AWS S3, GCS & Azure Blob ingestion alongside bi-directional Supabase PostgreSQL sync with Row Level Security.

Sentinel — Autonomous Multi-Agent PR Reviewer ([GitHub](#))

Python, FastAPI, LangGraph, Python AST, Pydantic v2, GitHub API

- Multi-Agent State Graph:** Engineered a cyclical LangGraph state machine dispatching PR diffs concurrently across Security (OWASP/CWE rules), Performance (N+1 queries/async stalls), and Architecture specialist agents.
- AST Context & Consensus Core:** Built a Python AST analyzer mapping diff hunks to enclosing function and class scopes, paired with a severity-weighted deduplication synthesizer and Pydantic v2 schema guardrails.
- Hybrid Execution & Dispatch:** Integrated a Multi-LLM BYOK gateway with a **<20ms** zero-cost deterministic AST fallback engine, live GitHub PR fetch/publish endpoints, and a **19/19 passing test suite**.

Synapse — Enterprise Hybrid Search & RAG Engine ([GitHub](#))

Python, FastAPI, PostgreSQL (pgvector), BM25, Redis, Ragas

- 4-Stage Hybrid Retrieval:** Fused dense vector embeddings (pgvector HNSW) and BM25 lexical search via Reciprocal Rank Fusion (k=60) and Cross-Encoder re-ranking across a 6-domain corpus with table-preserving contextual chunking.
- Ragas Grounding & Citations:** Achieved **85.7% Context Precision** and **100% Context Recall** on enterprise benchmarks (outperforming 55.0% dense baselines), synthesizing clickable page-cited answers with 100% faithfulness.
- Redis Semantic Caching:** Engineered SHA-256 and cosine similarity (≥ 0.95) query deduplication serving cached responses in **<25ms** at **\$0.00** token cost, cutting LLM API expenses by **60%** (19/19 passing tests).

Chess: The Celestial Pawn — Mobile Chess Engine ([Play Store](#))

C#, Unity Engine, State Machines, Memory Profiling, Google Play

- Deterministic Game Architecture:** Engineered custom move-validation algorithms, bitboard-inspired state evaluations, and a decoupled finite state machine managing dual-mode board transitions.
- Runtime Optimization:** Profiled managed heap allocations and draw calls via Unity Profiler, optimizing sprite atlasing to eliminate GC spikes and sustain a locked **60 FPS** across low-tier Android hardware.

Block Glow — Native Android Puzzle Engine ([Play Store](#))

Kotlin, Android SDK, MVVM, Room DB, Coroutines, ProGuard/R8

- Custom 2D Rendering:** Developed a native Android puzzle engine with double-buffered Canvas rendering and object pooling, sustaining **60 FPS** while reducing runtime garbage collection pauses by **42%**.
- Persistence & Binary Shrinking:** Implemented Room DB and Coroutine state persistence (**<140ms** cold-boot latency) and configured ProGuard/R8 optimization to cut production APK size by **36%**.

EDUCATION & CERTIFICATIONS

Chandigarh University | Bachelor of Engineering in Computer Science and Engineering

Graduated: 2026

Certifications: Oracle Cloud Infrastructure Certified Architect Associate (Oracle, 2024) | Cyber Security Certified Associate (Solitaire Infosys, 2022)