SHIKHAR SRIVASTAVA

Senior Software Engineer | Backend & Distributed Systems

Cloud & AI Infrastructure • Payments & Transaction Processing • Real-Time Systems

Delhi, India | +91 7252915788
shikhar.cs.eng@gmail.com | Shikhar.dev | LinkedIn.com/in/ItisShikhar | GitHub.com/ItisShikhar

SUMMARY

Senior Software Engineer and builder with 8+ years building high-scale backend and distributed systems for gaming, payments, and AI products. At Junglee Games (Flutter Entertainment), owned Node.js/TypeScript services for millions of users and INR 5 crore+ (~$0.5M+) in daily transactions; cut matchmaking latency by ~50%. Led design reviews, mentored engineers, and shipped production systems.

Since 2023, solo founder-engineer: built and operated MaxReel, an AI video platform that served 500+ creators with asynchronous inference workflows sustaining 100+ concurrent jobs (since discontinued), and published open-source tools. Solely responsible for the architecture of GCP services, AI workflows and asynchronous processing, as well as for reliability and transactions.

Also building LLM-powered and multimodal AI systems, including hybrid retrieval and evaluation (RAG), AI pipelines and multimodal apps.

EXPERIENCE

Founder & Engineer | Crescent Game Studios *2023 - Present*

MaxReel - AI Video Generation Platform

- Built and shipped a cloud-native AI inference orchestration harness on GCP (Cloud Run, Firestore, Cloud Tasks) coordinating model calls, streaming delivery, and FFmpeg rendering for short-form video.
- Designed fault-tolerant async workflows (retries, backoff, idempotency, concurrency limits) sustaining 100+ concurrent render jobs.
- Cut end-to-end generation latency to 30s-2m (**~40% faster**) by parallelizing AI and rendering stages.

CaptionCat - Open-Source TypeScript Caption Engine | github.com/ItisShikhar/captioncat

- Published an MIT-licensed TypeScript package (@captioncat/caption-engine) with Node and browser APIs and a CLI that renders animated captions to video and images, built on an Entity Component System for reusable layout, animation and effect presets.
- Integrated FFmpeg, Skia Canvas and pluggable transcription adapters (OpenAI, ElevenLabs, Sarvam).

Ludojack - Real-Money Gaming Platform

- Built secure **wallet and payment services** for real-money deposits, withdrawals, gameplay transactions, and ACID-compliant reconciliation workflows using Cloud Tasks and Cloud Scheduler.
- Scaled Cloud Run backend services to **hundreds of thousands of daily requests** and implemented geo-fencing, device fingerprinting, and anti-fraud/anti-collusion controls

Software Development Engineer II (SDE-2) | Junglee Games (Flutter Entertainment) *2019 - 2023*

- Owned core Node.js/TypeScript backend microservices for gameplay state, tournaments, matchmaking, and financial ledger services serving **millions of active users**.
- Reworked matchmaking algorithms and real-time game-state synchronization, reducing polling overhead, cutting matchmaking latency by **~50%**, and improving player retention during peak traffic.

- Designed wallet, payment, and gameplay transaction services using REST APIs, PostgreSQL/NoSQL, and async processing; processed **INR 5 crore+ daily transaction volume** with ACID guarantees across deposits, withdrawals, gameplay, and settlements.
- Integrated Apache Kafka event streams for asynchronous gameplay events, audit logging, and decoupled tournament state propagation across microservices.
- Built a bot framework that increased matchmaking fill rates by **80%**.
- Improved production reliability by reducing crash rates by **25%** through profiling across Unity, Flutter, and Node.js microservices
- Mentored 2+ Unity engineers in Node.js and Flutter, cross-skilling team members into full-stack contributors and leading design reviews across client-server boundaries.

Software Development Engineer | Suventure (VirtuGO, Denmark) *2018 - 2019*

- Built core gameplay and real-time simulation features in Unity for VirtuGO, a virtual cycling training platform for Olympic athletes, optimizing physics and rendering performance while working with engineering teams across India and Europe.

SELECTED OPEN-SOURCE PROJECTS

Gaming Regulations RAG | github.com/ItisShikhar/gaming-regulations-rag

- Built a jurisdiction-aware hybrid retrieval system (Qdrant dense vector search, BM25 with reciprocal rank fusion, metadata filtering) over ~7K clause-level chunks from 100+ India Code records, the UK Gambling Commission, the New Jersey Division of Gaming Enforcement, and Gaming Laboratories International standards. On a 48-question hand-labeled dev set, filtered hybrid reached **~95% Recall@5** vs ~80-83% for dense-only, BM25-only and unfiltered hybrid.

gg-friggin-ez - Multilingual Toxicity & Profanity Screener | github.com/ItisShikhar/gg-friggin-ez

- Built an open-source moderation engine for code-mixed Indic game chat, powered by JEV (System 1 decision model). 100% evasion recall and 0% false bans at 430ms p95 latency (in a small sample validation test).

EDUCATION

Bachelor of Technology - Computer Science | *2014 - 2018*
GLA University, Mathura, India

SKILLS

- **Backend & Distributed Systems:** Node.js, TypeScript, REST APIs, Microservices, Event-Driven Architecture, Apache Kafka, Async Processing, Concurrency, Distributed Transactions, Caching, Retries, Backoff, Idempotency, Video Streaming
- **Cloud & Infrastructure:** Google Cloud Platform (Cloud Run, Firestore, Cloud Tasks, Cloud Scheduler, Pub/Sub), Microsoft Azure (App Service, Functions, Service Bus, Storage, SQL), Docker, CI/CD, Serverless
- **AI Product Engineering:** LLM integration, async inference orchestration, model routing, hybrid retrieval (dense + BM25, RRF), retrieval evaluation (Recall@k, MRR), PyTorch (working knowledge), multi-modal apps, MCP, LoRA
- **Data & Languages:** PostgreSQL, NoSQL, Firestore, MongoDB, Realtime Database, TypeScript, JavaScript, C#, Dart; Python (working knowledge)
- **Tools:** FFmpeg, Unity, Flutter, Git, GitHub Actions
- **AI-assisted Development:** Claude Code, Cursor; spec-first workflow, test-driven review of generated code