Email : hello@aryank.space
https://aryank.space Mobile : +91-8421-911-353

Aryan Kathawale

linkedin.com/in/kiritocode1 | github.com/kiritocode1

Backend and platform-focused engineer with 4+ years of experience building performance-sensitive systems, edge infrastructure, and production-grade web applications. Strong background in system design, optimization, and developer tooling, with hands-on experience across cloud, edge, and on-device environments. Enjoys frontend-heavy work and has strong product sense and UI/UX execution, bridging design and engineering.

Technical Skills

Languages: TypeScript, Go, Python, Rust, C, SQL
Backend & Platform: Node.js, Hono, Fastify, Express, Serverless Architectures, Edge Computing
Frontend: React, Next.js, React Native, Expo
Cloud & Infrastructure: AWS (Lambda, API Gateway, S3), Cloudflare Workers, Docker
Data & Messaging: PostgreSQL, MongoDB, Redis, Kafka, RabbitMQ
AI / ML Systems: ONNX Runtime, WebGPU, TensorFlow, PyTorch, LLM Integration
Tooling & DevEx: Git, pnpm, Bun, Deno, Tailwind CSS, Figma

Experience

Senior Full-Stack Engineer Sept 2025 – Present
Dreamcare Developers LLC *Pune, India*

- Led a re-architecture of the frontend and backend stack into a **pnpm-based monorepo**, enabling shared components across multiple government-facing applications, including sites for the Maharashtra Police Department.
- Designed and built an internal component registry inspired by **shadcn/ui**, enforcing **UX4G** accessibility and performance guidelines across teams.
- Reworked backend services using **Hono** and **Convex** to better support evolving requirements from government intelligence and public-sector stakeholders.
- Reduced end-to-end development complexity and review overhead, cutting average pull request turnaround time by **50%**.

Full-Stack Engineer Aug 2022 – Sept 2025
AxiaNext LLC *Kyoto, Japan (Remote)*

- Designed and owned a production-grade web data extraction platform in **TypeScript** using **Puppeteer**, supporting multiple configurable sources with retry and recovery pipelines.
- Engineered fault-tolerant scraping workflows including proxy rotation, CAPTCHA handling, adaptive selectors, backoff strategies, and DOM diffing, achieving **90% extraction accuracy** across unstable third-party sites.
- Reduced manual research and data collection effort by **40%**, unblocking downstream analytics and operations teams.
- Acted as the primary engineer responsible for reliability, debugging, and rapid iteration as upstream sites frequently changed.
- Worked on authentication, data access, and API integration for a **React Native** mobile application consuming the same multi-tenant backend services.
- Participated in production testing, monitoring, and iterative rollouts, debugging issues across backend services, cloud infrastructure, and mobile clients.

Backend Engineer (Intern) May 2021 – Dec 2021
Yottalligent *Mumbai / Remote, India*

- Worked on backend services for a multi-tenant e-commerce platform, supporting isolated tenants, shared infrastructure, and configurable business logic.
- Contributed to migrating production APIs from **GraphQL** to **JSON-RPC**, simplifying request flows, improving runtime reliability, and reducing client–server coupling.
- Built and optimized asynchronous **Node.js** workflows and service initialization paths, improving backend performance and developer iteration speed.
- Refactored selected performance-critical backend components into **Go**, reducing compilation times by **25%** for those services.
- Deployed and operated backend services on **AWS** using a serverless architecture (**Lambda**, **API Gateway**, **S3**), gaining hands-on experience with production cloud infrastructure.

Projects

Voices - On-Device TTS | *ONNX Runtime, WebGPU, TypeScript/JavaScript, Custom Quantized Models* 2025

* Developed a lightning-fast, fully on-device text-to-speech system running locally with zero cloud dependency, ensuring complete privacy and no API calls.
* Achieved extreme performance via ONNX Runtime optimizations and quantization: 12,164 tokens/s on RTX 4090, 2,509 on M4 Pro WebGPU, 1,263 on M4 Pro CPU.
* Outperformed existing solutions (Flash v2.5: 287, TTS-1: 82, Gemini 2.5 Flash TTS: 24) in low-latency, edge inference scenarios.
* Targeted real-time, privacy-first TTS for consumer devices (desktops, laptops, mobiles) with minimal computational overhead.

blank-lb: Cloudflare Workers Load Balancer | *Cloudflare Workers, TypeScript, Effect-TS* 2025

* Created a type-safe, composable load balancer library running entirely on Cloudflare's edge network, avoiding reliance on vendor-specific enterprise features.
* Implemented advanced routing strategies: round-robin, geo-steering (continent/country/region/colo), health-based, failover with multiple recovery modes (fail-forward, async-block, promise-any).
* Added observability via custom response headers (X-Load-Balancer-*) for endpoint, latency, retry count, plus support for canary releases, blue-green deployments, and sticky sessions.
* Engineered routing and failover logic to operate within Cloudflare Workers' CPU and cold-start constraints, favoring fast decision paths and minimal per-request overhead.

Better PDF Reader | *Next.js, React, TypeScript, shadcn/ui, Biome, Indexed-DB* 2026

* Built a performance-focused PDF reading platform with client-side analytics capturing reading time, session behavior, and interaction patterns.
* Designed a lazy-loaded PDF rendering and indexing pipeline that initializes in under 100ms regardless of document size, enabling fast first interaction and predictable client-side performance.

Education

Sinhgad College Of Engineering Pune, Maharashtra

Bachelor of Engineering in Computer Science, Major in Machine Learning, ***7.64 GPA***