

SHAMIL DILSHAN PREMATUNGA

SENIOR SOFTWARE ENGINEER | MSC IN DS & AI (READING)

No 86, The Peak, Horana, Sri Lanka | +94 70 256 1890 | shamildilshan.b@gmail.com | [LinkedIn](#)

SUMMARY

Senior Software Engineer with 6+ years of experience designing and building scalable cloud-native and enterprise platforms. Experienced in backend engineering, distributed systems, Kubernetes, AWS, and full-stack development using Go, TypeScript, React, Python, gRPC, PostgreSQL, and DynamoDB. Currently leading engineering initiatives at H2O.ai across architecture, technical ownership, team leadership, and product delivery. Hands-on experience building AI-powered services using LLM orchestration, RAG, vector search, streaming APIs, and cloud-native deployment.

TECHNICAL SKILLS

GoLang	Python	React/TS	AWS	Docker + K8s
gRPC + REST	SSE Streaming	LLM + RAG	Vector Search	OIDC + JWT
Postgres	DynamoDB	Athena DB	Helm	GitHub Actions

PROFESSIONAL EXPERIENCE

Senior Software Engineer, H2O.ai **Feb 2025 - Present**

- Owning and leading mission-critical cloud-native platforms at H2O.ai, driving architecture decisions, technical execution, and delivery across enterprise products while mentoring engineers and contributing hands-on across the stack.
- Building AI-powered platform capabilities using Go, REST/SSE streaming APIs, multi-provider LLM integration, RAG pipelines, vector search, automated documentation ingestion, and OIDC/JWT authentication.
- Hands-on with LLM orchestration, prompt engineering, context-window optimization, and deploying production services on AWS using Docker, Kubernetes, Helm, and CI/CD tooling.
- Collaborating with stakeholders to align engineering work with business goals, improve platform scalability, and deliver secure, production-grade systems.

Software Engineer, H2O.ai **Mar 2022 - Feb 2025**

- Worked on building scalable web apps using Python, H2O Wave, React, TypeScript, GoLang, gRPC, Postgres, DynamoDB, Athena and AWS. Mentored interns, supported sprint planning, and gained practical experience with Jenkins pipelines.

Research Engineer, Dialog - UoM Laboratory **Jul 2020 - Mar 2022**

- Worked on opensource 5G Cellular Network Implementation Project.

Software Engineer, Applova Inc. **Feb 2020 - Jul 2020**

- Worked on restaurant based web application.

PROJECTS

Forward Deployed Engineer (Consulting) - Plantation SaaS

Sep 2026 - Present

- Consulting on a SaaS solution targeted at the plantation sector, working closely with stakeholders to understand operational needs and translate them into practical product features and workflows.
- Acting as a Forward Deployed Engineer across solution design, implementation support, integrations, deployment coordination, and production problem solving.

Platform Assistant

Jul 2026 - Present

- AI-powered code generation service that translates natural language into executable code and configuration across multiple products and languages. Built end to end with streaming LLM responses, RAG over product documentation, vector retrieval, authentication, and Kubernetes deployment.
- Tech Stack: GoLang, REST + SSE, multi-provider LLMs, RAG, vector search, automated doc scraping, OIDC/JWT, Docker + K8s + Helm, AWS

Admin Center

Mar 2022 - Present

- Production grade, cloud native platform developed for managing and administrating environments covering user management, security, billing tracking, observability, and identity management across enterprise deployments.
- Tech Stack: GoLang, TS, gRPC + REST, Protocol Buffers, PostgreSQL (AWS RDS), Docker + K8s, Github Actions

Insights Dashboard

April 2024 - Present

- Enterprise grade internal analytics platform. It tracks telemetry analytics across all environments.
- Tech Stack: GoLang, TS, Python, AWS Athena, DynamoDB, S3, Docker + K8s + Helm, Github Actions

Access Manager

April 2025 - Present

- Internal access control platform built from the scratch. It lets team members request, approve, and track time limited access to EKS clusters and AWS accounts with role based permissions, real-time notifications, and a full audit trail. This is a production grade internal tool solving a real infrastructure security problem.
- Tech Stack: GoLang, gRPC web, Protocol Buffer, AWS SDK, Slack Webhooks, TypeScript, MUI, Docker, AWS DynamoDB, SQS, Github Actions.

HIGHER EDUCATION

M.Sc. in Data Science and Artificial Intelligence (Reading)

Jan 2026 - Present

University of Moratuwa, Sri Lanka

Research Topic: Understanding Security Vulnerabilities in LLM-Generated Web Application Code: A Red Teaming Approach to Characterising the Functional-Security Correctness Gap.

B.Sc. (Hons). Electronic and Telecommunication Engineering

Nov 2015 - Jan 2020

University of Moratuwa, Sri Lanka

- GPA: 3.55/4.2 - Second Class Upper Division
- FYP: Safety enhancement of pedestrian crossings (Final year project)

ADDITIONAL INFORMATION

Sports: Table Tennis, Chess

Youtube: [Tekcerpts YT](#)

Certifications: [LinkedIn Certificates](#)

Blogs: [Medium Blogs](#)

REFERENCES

Prof. (Mrs.) Dileeka Dias

Deputy Vice Chancellor
University of Moratuwa, Sri Lanka
dileeka@uom.lk
+94 777 688 861

Mr. Vijitha Ekanayake

Manager Fullstack Development
vijitha.ekanayake@h2o.ai
+94 777 247 339