MAHMOOD UL ISLAM

Software Engineer | Cloud & AI Specialist

+1-306-261-8482 | mahmood.islam@globalcrowndigital.com | portfolio-mahmoodulislam.vercel.app | github.com/MahmoodUlislam | linkedin.com/in/mahmoodislam

AI-focused software engineer with 8+ years shipping production systems end-to-end and a decade of entrepreneurship — citation-grounded RAG, multi-agent LLM systems, vision models, and the AWS infrastructure beneath them.

AREAS OF EXPERTISE

Programming Languages: Python, TypeScript/JavaScript, Rust, Swift, Kotlin, C (embedded/ESP-IDF), SQL

Frontend Development: React.js, Next.js, Vue.js/Quasar, React Native (Expo), SwiftUI

Backend Development: Node.js, Express.js, FastAPI, Next.js APIs, REST, GraphQL; WebSockets, Socket.IO, SSE, WebRTC, Chime SDK

Databases & Storage: PostgreSQL, MySQL, Amazon RDS, DynamoDB, MongoDB, SQLite, Qdrant (vector); Prisma; S3

Cloud & DevOps (AWS): ECS Fargate, Lambda, EC2, API Gateway, VPC/ALB/NLB, EventBridge, Step Functions, SNS/SQS/SES, Cognito, IAM, Secrets Manager, CloudWatch, IoT Core, CloudFront, Elastic Beanstalk; IaC: CDK, CloudFormation; GitHub Actions, Docker

Generative AI & Agents: Amazon Bedrock (Agents, Knowledge Bases, Guardrails), RAG with source-grounded citations (BGE-M3, Qdrant, re-ranking), multi-agent orchestration (LangGraph), self-hosted LLM serving (vLLM, speculative decoding, KV/prefix caching), prompt-injection defence, LLM-as-judge evaluation

ML & Computer Vision: PyTorch, Hugging Face Transformers, ViT/DINOv3 transfer learning, SageMaker (training → scale-to-zero inference, MLOps), semantic segmentation, Amazon Rekognition, AprilTag calibration, Pandas, NumPy, Scikit-learn

NLP, Speech & Medical AI: Amazon Comprehend (Medical), Transcribe, Polly; ASR (Parakeet, Whisper), NeMo diarization, GLiNER/spaCy NER, OCR (Docling, Tesseract)

Data Engineering: AWS Glue, Kinesis, Data Firehose, Athena, Redshift

Embedded & IoT: ESP-IDF (ESP32-S3/C5/C6), WiFi CSI sensing, 60 GHz mmWave radar, AWS IoT Core (MQTT/TLS)

Security & Identity: OAuth 2.0/OIDC, Cognito, Auth0, RBAC, TOTP MFA, YubiHSM, tamper-evident audit logging

EXPERIENCE

Case Intelligence Assistant Inc.

Founding Engineer **March 2026 – Present**

Location: Remote, Canada

Project: "CiA — Case Intelligence Assistant" (caseintelligenceassistant.com) — **air-gapped AI investigative file-analysis platform for law enforcement**, fully offline on an NVIDIA DGX Spark (128 GB unified memory)

Technologies: Python 3.12, FastAPI, SQLAlchemy, LangGraph, vLLM, Gemma 4 31B, BGE-M3, Qdrant, Parakeet/Whisper ASR, NeMo, GLiNER, Docling, Tesseract, SQLite, YubiHSM 2, CUDA

- Sole engineer of the entire platform — FastAPI backend, SPA frontend, vector search, security, and deployment — from first commit to CEO-signed **V1.0 in under 3 months**.
- Built a production **RAG pipeline where every answer cites its exact source** (page, video frame, transcript timestamp): BGE-M3 → Qdrant ACORN retrieval → cross-encoder re-rank → Gemma 4 31B on vLLM streamed over SSE; retrieved evidence treated as untrusted to defend against prompt injection.
- Cut large-case, cross-document inconsistency analysis from **~6 hours to ~1 hour** — the kind of review that costs an investigator a full day — through self-hosted inference tuning (vLLM speculative decoding, KV/prefix caching) and entity-graph pruning.
- Multimodal ingestion of 13 document formats plus audio/video: Docling/Tesseract OCR, Parakeet ASR with Whisper fallback, NeMo diarization & voice re-ID, zero-shot NER (GLiNER); crash-resumable LangGraph pipelines.
- Court-defensible by design: RBAC + TOTP MFA, YubiHSM 2 key custody, hash-chained tamper-evident audit log, Crown-disclosure export (R v. Stinchcombe), visible AI-generated labels, EN/FR bilingual UI.

Global Crown Digital Corporation

Director & Founder **August 2024 – Present**

Location: Canada

Project: "BlindSense" (March 2026 – Present) — camera-free elderly wellness monitoring fusing WiFi CSI sensing with 60 GHz mmWave radar — no cameras, wearables, or microphones

Technologies: C (ESP-IDF v6, ESP32-S3/C5/C6), Rust, TypeScript, Next.js, SwiftUI, AWS IoT Core, DynamoDB, Bedrock, CDK, APNs

- Founded and operate a Canadian software corporation delivering AI, cloud, and IoT product engineering; conceived and built BlindSense end-to-end as sole engineer.
- ESP-IDF C firmware across a 3-node ESP32 kit — real-time WiFi-CSI capture and multi-sensor fusion at the edge, OTA A/B updates, per-device X.509 — all safety-critical decisions made on-device, never in the cloud.
- Event-driven AWS backend in CDK: IoT Core MQTT → Rust Lambda services → DynamoDB/SNS/EventBridge; a deterministic stats engine decides, Amazon Bedrock (Claude) only narrates pre-computed flags with claim-lint guardrails.
- Shipped a Next.js real-time dashboard (SigV4-signed MQTT over WebSocket, 3D radar point cloud) and a SwiftUI iOS app with APNs push signed ES256 over HTTP/2 from Rust.

Virtual Health Hub

Senior Software Engineer | Cloud & AI Specialist — **March 2025 – August 2026**

Software Developer — **August 2024 – February 2025**

Location: Saskatoon, Saskatchewan, Canada

Project: "LifeLine" (lifelineapp.ca) — virtual mental-health treatment platform

Technologies: React Native (Expo), Next.js, Vue.js/Quasar, Node.js, Python, PostgreSQL/Prisma, Socket.IO; AWS: Bedrock Agents, Chime SDK, Transcribe, Polly, ECS Fargate, Cognito, Kinesis

- Managed the project end-to-end — development, cloud infrastructure, and solution architecture — building a multi-agent AI system on Amazon Bedrock (chatbot, medical-scribe, and risk-analytics agents) with RAG, voice via Polly/Transcribe, and Chime SDK video with ambient-listening summarization; HIPAA-aligned VPC.
- Built appointment booking and medication management with analytics; developed native Swift/Kotlin Expo modules bridging Chime SDK — published as react-native-chime-sdk, an open-source TurboModule/Fabric library used in the production app.
- Integrated Apple HealthKit / Health Connect (30+ biometrics auto-syncing to the care team) and delivered clinician portals in Vue/Quasar + Next.js.
- **First author** of a REB-approved safety evaluation of the platform's agentic AI — 4,500+ scored interactions across two datasets, validated against expert clinical review — manuscript prepared for *Mayo Clinic Proceedings: Digital Health*.

Project: "SkinScan" (skinscanapp.ca) — AI wound-care imaging platform, end-to-end

Technologies: Swift (iOS), Next.js, Node/Express, PyTorch, Hugging Face, AprilTags; AWS: SageMaker, Bedrock, Rekognition, Comprehend Medical

- Trained a unified DINOv3 ViT-L/16 for 10-class wound classification + pixel-level segmentation in one forward pass; curated an 11-source clinical dataset to eliminate severe class imbalance; deployed on SageMaker async endpoints with scale-to-zero autoscaling; AprilTag calibration for cm^2-accurate measurements; 3 Bedrock agents for staging and recommendations.
- Built the delivery surface: Swift iOS capture app with AR guidance and Rekognition pre-screening, Next.js clinician portal, Node/Express backend with Cognito RBAC, SSE streaming, and Comprehend Medical (ICD-10-CM, RxNorm, PHI).

Project: "VHH Command Center" — telehealth fleet-operations portal

Technologies: Next.js, Socket.IO, WebRTC, MapLibre, Drizzle, SwiftUI; AWS: CDK, Elastic Beanstalk, CloudFront, Cognito

- Shipped a production control-room portal monitoring telehealth deployments across 11 northern Saskatchewan communities: live geospatial fleet map, machine telemetry with alert acknowledge/resolve workflow, and per-community topic channels (operations/clinical/maintenance) with real-time chat.
- Built phone-style WebRTC voice/video ring-out to on-site assistants — web peers and a native SwiftUI iOS companion (TestFlight beta) — plus Web Push critical alerts and a pluggable integration-adapter seam for third-party machine data.

Saaska Software Inc.

Software Engineer — **April 2023 – June 2024**

Location: Saskatoon, Saskatchewan, Canada

Project: "esiKidz" (esikidz.com) — childcare management software

Technologies: JavaScript, React.js, React Native, Node.js, REST APIs, MySQL, AWS (Lambda, S3, API Gateway)

- Developed enterprise SaaS with React.js & React Native apps communicating via Node.js API services; built Python web-crawling and data pipelines (BeautifulSoup, Pandas, NumPy).
- Automated the deployment toolchain for instant provisioning of new services; built internal apps and custom libraries for client management; improved reliability, scalability, performance, and security.

SELISE Digital Platforms

Software Engineer — **June 2022 – August 2022**

Project: "Sunrise-club" (sunrise.ch/en/moments) — event booking site for Sunrise (Swiss telco) users

Technologies: TypeScript, React.js, Next.js, MongoDB, REST APIs, OAuth 2.0/OIDC, Storyblok CMS

- Maintained WAF-fronted URL redirection on the production site; built Next.js APIs & middleware with OAuth 2.0 + OIDC; implemented Storyblok CMS and MongoDB.
- Maintained 3 test environments, codebase documentation and standards, and built a complete language-translation system.

Quantic Dynamics Ltd.

Web Developer — **February 2019 – May 2022**

- Built fully responsive web applications using React.js, Next.js, and Node.js with MongoDB and Firebase.

EDUCATION & CERTIFICATIONS

- Applied Certificate in Website Design and Development — Saskatchewan Polytechnic
- Bachelor of Science — Ahsanullah University of Science & Technology (WES Canadian equivalency: 4-year bachelor)
- **NVIDIA DLI — Generative AI with Diffusion Models** (June 2026), ID: tZ_6rIcjRReMksD-FZbAOw
- **NVIDIA DLI — Fundamentals of Accelerated Computing with CUDA Python** (April 2026), ID: ZVO5ZmCKQY-YfG2iC6wfzw
- **AWS Certified Machine Learning Engineer – Associate** — Validation No. 778b21fc058d42a39b13860d3b4afec8