Ibrar Yunus

FULL-STACK AI ENGINEER · RESEARCH · DEVELOPMENT · DEPLOYMENT

Email solutions@ibraryunus.com · **Phone** +44 7393 741 027 · **Location** Tokyo, Japan
Web ibraryunus.com · **LinkedIn** linkedin.com/in/ibrar-yunus

PROFESSIONAL SUMMARY

British Full-Stack AI Engineer building production-ready AI products end-to-end. From RAG pipelines and LLM grounding to scalable Next.js + Cloud Run deployments — I take research to revenue. Approximately 10 years across enterprise (THG, Gifftid), independent contracting and research roles; deployed AI serving 7M+ monthly users. Now based in Tokyo and committed to building my career here, seeking a full-time role with visa sponsorship (approx. 1–2 months to issue).

PROFESSIONAL EXPERIENCE

Contract Full-Stack AI Engineer & Data Scientist

Independent Contractor — sole trader (multiple clients) · Remote, overseas clients · **May 2025 — Present**

- **Loredana — AI Coaching SaaS:** architected a full-stack Next.js + Firebase platform; built "Sage Journal" with Gemini for sentiment-aware coaching feedback, real-time Recharts dashboards (6-pillar resilience radar, somatic body mapping), and role-based Firestore auth.
- **Telleosis / MyJamb — EdTech AI Tutor:** engineered a FastAPI backend on Cloud Run with a RAG pipeline over BigQuery Vector Search (cosine distance) to retrieve lesson contexts; designed SQL gamification for leaderboards and randomised exam generation.
- **Kotchakorn — Language-Learning AI:** interactive Thai-learning app in Python Shiny + BigQuery with Gemini 2.5 Flash streaming proficiency-tuned lessons; automated PDF and flashcard generation.

Lead Data Scientist

Gifftid Ltd. (all-in-one AI platform for SMEs) · London, UK · Remote · **Sep 2024 — May 2025**

- Architected and shipped the complete GenAI platform — GCP infrastructure, CI/CD for Machine Learning, containerised deployments (Docker / Kubernetes), monitoring, access controls, chatbots, vector databases, and analytics dashboards.
- Partnered with the CEO on strategic AI direction; identified market gaps and shipped SME-focused tooling that positioned Gifftid as a leading AI-driven SME platform.
- Built in a small core team alongside a former Google Ads engineering lead and an MIT PhD product lead, driving development through close technical debate as an equal contributor.

Data Scientist & Software Engineer

The Hut Group (THG) (global e-commerce, health & beauty) · Manchester, UK · **Oct 2018 — Sep 2024**

- Built production GPT chatbots grounded to company data with hallucination prevention, moderation guardrails, RAG, and chain-of-thought / tree-of-thought prompting (LangChain, NVIDIA NeMo Guardrails).
- Shipped multi-modal vision-language outfit recommendation chatbots using InstructBLIP and GPT models on Azure; deployed AI projects serving 7M+ monthly users.
- Built topic detection and sentiment-analysis pipelines using guided LDA, RoBERTa, BERT and DistilBERT; integrated Machine Learning services into production APIs and backend systems.
- Featured at THG Ingenuity Future of Commerce 2023 Industry Innovators Conference; trained and project-led other data scientists.

Research Internships & Final-Year Project

University of St Andrews · Trinity College Dublin · Maynooth University · Augary (US) · DFKI (Germany) · **2015 — 2018**

- Machine Learning on 2,500 Full-HD videos (12-node cluster) for video-encoding optimisation; ROS/TurtleBot 3D room-change detection extending MIT's Kintinuous SLAM; real-time on-device lane detection (patent-pending, IEEE-published); spatial 5.1 audio research with Google and SFI Ireland.

TECHNICAL SKILLS

AI / MACHINE LEARNING: LLMs · RAG · Generative AI · Machine Learning · NLP · Computer Vision · Multi-Modal · Fine-tuning · Prompt Engineering · Embeddings · Vector Search · Sentiment Analysis · Topic Modelling · Recommendation Systems · Classical ML · Time-Series Forecasting · MLOps · A/B Testing

MODELS: GPT-4/5 · Claude · Gemini · Llama · BERT · RoBERTa · DistilBERT · BLIP-2 · InstructBLIP · X-CLIP · CLIP

PROGRAMMING LANGUAGES: Python · TypeScript · JavaScript · SQL · Rust · C++ · C · Go · Java · Bash · HTML / CSS

FRAMEWORKS & LIBRARIES: Next.js · React · Node.js · Tailwind CSS · FastAPI · Flask · Tauri · Shiny · LangChain · LangGraph · LlamaIndex · Hugging Face · PyTorch · TensorFlow · Keras · scikit-learn · Transformers · spaCy · NLTK · Gensim · Pandas · NumPy

CLOUD & DEVOPS: GCP (Cloud Run, BigQuery, Vertex AI, Cloud Functions, Pub/Sub, Cloud Storage, IAM) · Firebase · Firestore · Azure · Docker · Kubernetes · CI/CD · GitHub Actions · Linux · Git · REST · gRPC · OpenAPI · Microservices · Serverless

LLM TOOLING: OpenAI · Anthropic Claude · Google AI · Hugging Face · LangChain · LangGraph · LlamaIndex · NeMo-Guardrails · DialogFlow · BigQuery Vector Search · Pinecone · Chroma · ONNX · Ollama

DATA & ANALYTICS: BigQuery · PostgreSQL · MySQL · ETL · Data Warehousing · Recharts · Plotly · Streamlit

EDUCATION

MSc Dependable Software Systems (DESEM) — Distinction

University of St Andrews (UK) & Maynooth University (Ireland) · Erasmus Mundus Scholarship · **Sep 2016 — Jun 2018**

Thesis: 3D room-change detection on a TurtleBot with an RGBD sensor, extending MIT and Maynooth's Kintinuous visual SLAM — 83% (2D) / 100% (3D) accuracy.

BSc Computer Science

NUST — School of Electrical Engineering & Computer Science (SEECS) · Islamabad, Pakistan · **Sep 2012 — Jun 2016**

CERTIFICATIONS

- Google Cloud Platform Fundamentals: Core Infrastructure (Coursera) — 2021
- NVIDIA — Fundamentals of Deep Learning for Multi-GPUs — 2021
- Advanced SQL for Data Scientists (LinkedIn Learning) — 2021
- Advanced SQL for Data Science: Time Series (LinkedIn) — 2021

AWARDS & RECOGNITION

- Erasmus Mundus Joint Masters Scholarship — €49,000 (European Union)
- Gold Medal — Best Adjudged Industry Project (Undergraduate)
- MSc DESEM — Distinction
- Featured at THG Ingenuity Future of Commerce 2023 Industry Innovators Conference

PATENTS & PUBLICATIONS

Real-Time Lane Departure Warning System on a Lower-Resource Platform

US Patent Pending · IEEE Publication

Real-time lane detection and tracking on commodity Android hardware. Computer-vision pipeline in Python with CPU-intensive routines in C++ via ctypes, ported to Java for Android deployment. Extensively validated against the Caltech-Lanes dataset.

AI PRODUCTS & CLIENT BUILDS

- **Glass Invoice** `Next.js 16 · React 19 · Firebase · Gemini` — AI invoicing platform for UK sole traders and limited companies — instant PDF invoices, a real-time revenue and client analytics dashboard, AI tax guidance, and a conversational agent that answers questions grounded in the user's own financial data.
- **LINA** (bilingual AI learning studio): Mandarin learning system with HSK-scaled lesson synthesis on any topic, a real-time conversational simulator with tone feedback, spaced-repetition flashcards generated from lessons, and a persistent learner archive.
- **Nijikai** `Tauri 2 · Rust · headless Claude · agentic` — Tokyo settai planner for whoever books the client evening. Rust supervises cancellable headless-Claude process groups, streaming structured venue extraction under hard constraints; runsheets carry kamiza seating, discreet settlement and per-line last trains.
- **Euro-Japan Compass** `Next.js 16 · TypeScript · Tailwind v4 · Firebase` — client site for a European market-support consultancy serving Japanese firms — English and Japanese as two fully prerendered language environments, not one bilingual page; statically exported, no server runtime. Live at euro-japan-compass.web.app.
- Also shipped: GoChurch, Hayaat (E2E-encrypted AI journal), Kura (macOS Finder replacement), Sakurako, NotAlone, LinkVibe, Underscore, Vibelink, For Social Life — more at ibraryunus.com/ai-engineer.

PUBLISHED NATIVE APPS — RUST + TAURI

Signed, notarised desktop software, built and shipped solo — downloads and write-ups at ibraryunus.com/ai-engineer.

- **Saya** `Tauri 2 · Rust · Adblock-Rust` `macOS` — keyboard-first Zen browser. Ad and tracker blocking compiled in Rust, Tor tabs with .onion resolution, zero telemetry.
- **Mist** `Tauri 2 · Rust · Svelte 5` `macOS` — technical-brutalist native IDE. Full git workflow, LSP completion across 25+ languages, embedded PTY terminal, split editor.
- **Aks** `Tauri 2 · Rust · WebGL2` `macOS` — GPU photo editor and Story maker. WebGL2 grading pipeline — curves, 8-band HSL, local masks, film LUTs; ~8 MB, offline.
- **Andhera** `Tauri 2 · FFmpeg` `Windows` `macOS` `Linux` — a quiet dark video and audio player. 40+ formats via bundled LGPL FFmpeg, subtitles, ambient glow, visualiser, colour grading.
- **Naqsh** `Tauri · xterm.js` `Windows` `macOS` `Linux` — a designer's terminal emulator. Tabs, splits and panes; full ANSI/VT truecolour; optional CRT scanlines and phosphor glow.
- **Rait** `Tauri 2 · Rust · GCal` `macOS` — timezone-rail calendar for nomads. Unlimited synchronised time zones with per-zone public holidays and Google Calendar sync.

LANGUAGES

English — Native · Japanese — Beginner (studying) · Urdu — Professional · Thai — Limited · Chinese — Elementary

AVAILABILITY

TARGET ROLES: AI Engineer · Machine Learning Engineer · MLOps Engineer · Data Engineer · Data Scientist · Backend Engineer · Full-Stack Engineer

Location: Tokyo, Japan (JST) · Status: available on visa issue (approx. 1–2 months) · Seeking: full-time, permanent · Work style: on-site, hybrid or remote · Visa: open to sponsorship · Output: production, not proof-of-concept.