

LUKAS LICON

Santa Clara County, CA

lukas@licons.com • (415) 763-9382 • [linkedin.com/in/lukas-licon](https://www.linkedin.com/in/lukas-licon) • github.com/lukaslicon • lukaslicon.github.io/Portfolio/

SUMMARY

Applied AI engineer focused on building and operating **compact LLM systems**. I have designed, evaluated, and shipped **on-device/offline-capable language and speech solutions** with **observability** and **robust prompt/runtime controls**. Focused on deadlines, delivering results, and learning with every chance I get.

SKILLS

- **Languages:** C++, C, C#, Python
- **ML/AI (Frameworks & Techniques):** PyTorch; Hugging Face Transformers; **PEFT/LoRA** fine-tuning (causal LM); tokenization & **chat templates**; datasets/Trainer API; **mixed precision (fp16)**; ONNX; learning LangChain
- **Engines/Platforms:** Unreal Engine 5, Unity, whisper.cpp, llama.cpp, piper
- **Systems/Focus:** Real-time performance, on-device/offline inference, observability & runtime controls
- **Game AI:** Behavior Trees, State Machines, Pathfinding (A*, Dijkstra), MCTS

EXPERIENCE

Founding Engineer — Glade (YC23) • San Francisco, CA • Oct 2025 – Present

- Lead development of commercial **UE5 and Unity plugins** integrating open-source LLM and speech tech for building NPCs.
- **Trained and evaluated** compact LLMs (Llama 3.2-1B) for **local use in game engines**, targeting low memory and fast startup.
- Translated user feedback into roadmap, prioritized work in Jira, and coordinated weekly stand-ups and tasking.

AI/C++ Engineer — Glade (YC23) • San Francisco, CA • Oct 2024 – Sep 2025

- Built **UE5 C++ plugins** integrating **whisper.cpp, llama.cpp, ONNX, and piper** for **offline, low-latency speech-to-speech** pipelines.
- Implemented modular system prompts and curated data to support **dynamic relationship-style NPC interactions**.
- Shipped **user-facing UI** and robust state transitions to ensure stable in-game experiences.

Application Developer — Bluu Kazi • National Harbor, MD • Jul 2024 – Oct 2024

- Contributed to an Android app using **Kotlin, Android Studio, Jira, GitLab**, and **Figma** within a small agile team.
- Assisted with task prioritization, performance considerations, and iterative delivery.

PROJECTS

Crabity (Steam) — Unity/C# [Link to steam](#)

- Sole developer of **Time Trials** mode with **daily procedurally generated maps**; tuned 2.5D movement and shipped release.

C++ Fractal Tile Renderer — Personal [Link to Project](#)

- Asynchronous fractal renderer with message-queue design using **std::async** and a **ThreadSafeQueue**.

Python Navmesh Pathfinding — Personal [Link to Project](#)

- Implemented **bidirectional A*** over image-derived navmeshes to compute feasible routes.

C++ Matrix Operations (Sparse) — Academic/Personal [Link to Project](#)

- Built a sparse-optimized matrix calculator to improve memory and compute efficiency.

EDUCATION

B.S., Computer Science: Computer Game Design — University of California, Santa Cruz Summa

Cum Laude; GPA: 4.0 • Graduated Jun 2024 • Certification: Atlassian University — Jira Fundamentals (Jul 2024)