FRONTIER

Founding Staff [or Senior] Voice AI Systems Engineer

Company Overview |

Frontier is on a mission to make hundreds of millions of critical frontline workers superhuman through a hardware and software AI voice wearable. The company, which is headquartered in Chicago and in stealth mode, has multimillion dollar contracts with the U.S. Department of Defense and just raised $4.7M in funding from tier 1 investors, including the first investor in Oculus. The company has more than 10 of the most critical U.S. companies as paid pilot customers, including the largest U.S. steel company and two of the top five U.S. airlines. It has an extremely talented and ambitious team committed to building one of the most foundational companies of the 21st century.

Job Description |

Frontier is looking for a Founding Voice AI Systems Engineer (Senior or Staff level) who wants to do the most important work of their life. The selected individual will own huge parts of Frontier's voice platform (that lives on the cloud and on HW and their work will have extreme impact on the productivity, safety, and comfort of the critical frontline workers and warfighters powering and protecting the world. Being scrappy, resourceful, and flexible is critical to this role and it is 100% in-person in Chicago. Expected compensation is $200K-300K base, generous early-stage equity, relocation support, and good benefits. This role is ideal for someone who wants to (1) be a cornerstone member of a rapidly growing team and (2) iterate quickly on the cutting edge of AI voice interfaces (work deployed in the field in less than 1 year).

Responsibilities |

- Build and iterate low latency conversational voice pipeline, from raw audio capture through STT, retrieval augmented prompting, wake words, cloud LLM calls, TTS streaming, and back to playback, meeting tight latency and high accuracy requirements.
- Configure and tune orchestration within a pipecat-based stack (prompts/guardrails, component choices, dialog control), prioritizing evaluation over bespoke framework work.
- Ingest and govern enterprise data for LLM use so retrieval is scoped, secure, and personalized.
- Deliver accurate, traceable RAG by designing retrieval, grounding, and verification so answers cite sources and improve over time.
- Design, secure, and operate the cloud back end for a WebRTC-first, cloud voice platform; own the core APIs, auth, observability, scale, and reliability.
- Prototype lightweight services and tools as needed to unblock teams and keep the platform moving.
- Forward-deploy with customers to surface edge cases (pronunciations, jargon, accents) and harden real-world performance.
- Stay at absolute cutting-edge of voice AI and RAG-like systems.
- Collaborate across Android, back-end, firmware, and hardware to ensure smooth handoffs; document architecture and mentor as the platform scales.

2155 W Hubbard St, FL 1, Chicago, IL 60612
pmoeckel@frontieraudio.com | 860-305-5521

FRONTIER

Qualifications |

Required

- Hands-on experience (≥ 3 yrs overall, ≥ 1 yrs with modern LLM or end-to-end neural speech stacks) delivering production-grade conversational systems such as GPT-4/Claude-powered assistants, RAG chat-with-docs platforms, or streaming ASR or TTS programs (not Amazon-Alexa-Esque systems).
- Proficiency in Python and one systems language (Go/Rust/C++), plus comfort with gRPC/protobuf, REST, and event-driven messaging (e.g., Kafka or SNS/SQS).
- In-depth experience in prompt engineering, retrieval augmented generation, and vector databases (FAISS, pgvector, Milvus, Pinecone, etc.).
- Experience with streaming real-time audio in the cloud (WebRTC or equivalent low-latency) and tuning for latency, reliability, and cost.
- Experience prototyping lightweight auth/registry/telemetry/admin endpoints and full end-to-end voice flows with minimal back-end support.
- Must have ability to be qualified to work in the U.S. in less than 3 months and have no restrictions for international travel.
- Ability to commute to Frontier's Chicago office daily (relocation support in offer).

Preferred

- Experience shipping real-time conversational voice products (next-gen assistants, in-car AI copilots, field-service wearables, etc.).
- Experience deploying containerized services on Kubernetes/EKS/Fargate and automating infrastructure with Terraform or CDK.
- Experience measuring and optimizing end-to-end latency, power, and cloud cost for real-time voice or LLM pipelines.
- Solid grounding in OAuth/OIDC, JWT scopes, and secure key management for device-to-cloud authentication.
- Full-stack familiarity: ability to extend or hand off early back-end code into production-grade portals (PostgreSQL schema design, React/Next.js or similar UI, GraphQL or gRPC API gateways).
- Prior work integrating third-party SaaS platforms via webhooks or streaming adapters.
- Knowledge of ML observability and evaluation frameworks (Weights & Biases, PromptLayer, Evidently AI) and data-centric improvement loops.
- Contributions to open-source LLM or speech orchestration stacks (e.g. Pipecat, LangChain, LlamaIndex, Semantic Kernel, ESPnet, Vosk).

Frontier Audio Labs evaluates qualified applicants without regard to race, color, religion, sex, sexual orientation, gender identity, genetic information, national origin, age, disability, veteran status, or any other legally protected characteristics.

All inquiries should be directed to pmoeckel@frontieraudio.com with your portfolio and resume.

2155 W Hubbard St, FL 1, Chicago, IL 60612
pmoeckel@frontieraudio.com | 860-305-5521