

End-to-End Development of Speech-to-Text Systems for Voice Assistants in Distributed Web Frameworks

Sri Rama Chandra Charan Teja Tadi

Lead Software Developer, Austin, Texas.

Abstract: *Speech-to-text for voice assistants requires architectural accuracy in data flow, model behavior, and distributed coordination. This work formalizes the architecture of modular pipelines in web-based environments, where voice inputs are consumed as continuous streams, converted to spectral representations, and processed in stateless service interfaces. The session state is explicitly modeled to ensure contextual continuity, and consistency guarantees stable transcription across nodes. Security policies define stream-level confidentiality, and API contracts allow loosely coupled microservice calls. Modeling the end-to-end workflow from input to output and injecting observability throughout the stack, the system provides trustworthy, scalable voice interaction at the scale of production.*

Keywords: Speech-to-Text, Voice Assistants, Distributed Systems, Stateless Services, Session Modeling, API Contracts, Stream Processing, System Observability