

Sparrow AI

Technical Implementation Document

AI Partner Catalyst Hackathon - ElevenLabs Challenge

Executive Summary

Sparrow AI is a voice-first sales training platform that combines **ElevenLabs Conversational AI** with **Google Gemini 2.0 Flash** to create realistic AI sales prospects. Users practice cold calls, discovery conversations, and objection handling through natural voice interactions, receiving instant AI-powered feedback.

Core Integration: ElevenLabs (Voice AI) + Google Gemini (Intelligence) + Groq (Real-time Scoring)

1 System Architecture

1.1 Technology Stack

Layer	Technology	Purpose
Voice AI	ElevenLabs Conversational AI	Real-time voice conversations
Intelligence	Google Gemini 2.0 Flash	Persona generation, call analysis
Fast Inference	Groq (Llama 3.1 70B)	Real-time scoring during calls
Frontend	Next.js 15 + React 19	Modern web application
Database	Supabase (PostgreSQL)	Data persistence
Auth	Clerk	User authentication
Hosting	Vercel	Edge deployment

1.2 Data Flow

- Persona Generation:** User selects parameters → Gemini generates realistic prospect
- Voice Session:** ElevenLabs creates WebSocket connection with dynamic variables
- Conversation:** User speaks → ElevenLabs STT → LLM response → ElevenLabs TTS
- Analysis:** Call ends → Transcript sent to Gemini for detailed feedback
- Scoring:** Groq provides instant scores; Gemini provides deep analysis

2 ElevenLabs Integration

2.1 Conversational AI Implementation

We use the ElevenLabs React SDK (`@elevenlabs/react`) for seamless voice integration:

- **Signed URL Authentication:** Server-side URL generation for secure sessions
- **Dynamic Variables:** 15+ variables injected per session (persona details, call type, objections)
- **Voice Override:** Gender-appropriate voice selection (Jessica/Hope for female, Stokes for male)
- **Real-time Transcript:** Live transcript streaming during conversations

2.2 Key Features Used

ElevenLabs Feature	Application in Sparrow
Conversational AI Agents	Core voice interaction engine
Dynamic Variables	Persona customization per session
Voice Library	Multiple prospect personalities
Client-side Overrides	Voice selection based on gender
WebSocket Streaming	Real-time transcript display

3 Google Cloud / Gemini Integration

3.1 Gemini 2.0 Flash Usage

1. Persona Generation

- Input: Industry, role, personality type, difficulty level
- Output: Complete prospect profile (name, background, pain points, objections)
- Model: `gemini-2.0-flash-exp`

2. Call Analysis

- Input: Full conversation transcript + persona context
- Output: Detailed scoring across 5 dimensions with timestamped feedback
- Structured JSON output for consistent parsing

3. Coaching Suggestions

- Identifies missed opportunities with specific alternatives
- Highlights strong moments for reinforcement
- Provides actionable improvement tips

3.2 API Integration

```
// Gemini client initialization
import { GoogleGenerativeAI } from '@google/generative-ai';

const genAI = new GoogleGenerativeAI(process.env.GOOGLE_AI_API_KEY);
const model = genAI.getGenerativeModel({
  model: 'gemini-2.0-flash-exp'
});
```

4 Software Development Quality

4.1 Code Quality

- **TypeScript:** 100% type-safe codebase
- **Component Architecture:** Modular, reusable React components
- **API Design:** RESTful endpoints with proper error handling
- **Security:** Environment-based secrets, signed URLs, CORS policies

4.2 Performance Optimizations

- Edge runtime for low-latency API routes
- WebSocket connections for real-time voice streaming
- Groq for sub-second scoring responses
- Optimized bundle size with Next.js 15 App Router

5 Repository Structure

```
sparrow-ai/
src/
  app/                # Next.js App Router pages
  api/                # API routes
    calls/            # Call management
    personas/         # Gemini persona generation
    elevenlabs/       # ElevenLabs integration
  dashboard/          # Protected app pages
  components/         # React components
  lib/
    elevenlabs/       # ElevenLabs client & config
    gemini/           # Gemini client & prompts
    groq/             # Groq scoring client
  types/             # TypeScript definitions
docs/                 # Documentation
supabase/             # Database schema
```

6 Conclusion

Sparrow AI demonstrates deep integration of ElevenLabs Conversational AI with Google Gemini, creating a seamless voice-first experience. The technical implementation showcases:

- Quality software architecture with TypeScript and modern React
- Proper use of ElevenLabs voice AI capabilities
- Effective application of Gemini for intelligence and analysis
- Real-time performance with Groq inference
- Production-ready deployment on Vercel

Repository: <https://github.com/brn-mwai/sparrow-ai-partner-catalyst>
Live Demo: <https://sparrow-ai.brianmwai.com>