

音声認識アプリ開発レポート

🔗 プロジェクト構成

📱 フロントエンド1（HTML + JS,HTTPのシンプルページで検討）

- 録音機能付きのWeb UI（ボタン押下で録音開始 → 区切りで送信）
- Web Audio API による録音と無音検出（ScriptProcessorNode）
- 音声検出後は最大3バッファで録音し、録音が区切りで切り替わるごとに順番にサーバーへ送信する構成

区切りは 60秒 or 無音検出 + 1～5秒のバッファ

- 音声をFormDataとしてfetchでCloud Runの/webhookへ送信
- 結果テキストはチャット形式で画面に表示
- 音声波形表示（録音状態確認）

☁️ バックエンド1（Node.js on Cloud Run）

- Express + Multer による音声受信
- ffmpegで **FLAC (16kHz, mono)** へ変換
- GCS にアップロード : 一時保存とSTT変換効率UPのため
- Google Speech-to-Text API で文字起こし
- 結果をJSONでフロントに返却（CORS対応済）