Voice Assistant Terminal Commands:

**Tauri:**

Run server: npm run tauri dev

**Docker:**

Build: docker compose build --no-cache tts

Start/restart: docker compose up -d --force-recreate tts

Logging: docker compose logs -f tts  
  
**Sesame CSM Model Info:**

Backbone: llama-1B, Dim: 2048

Decoder: llama-100M, Dim: 1024

Audio tokenizer: Mimi/Encodec

Text Embeddings: 128256 -> 2048

Audio Embeddings: 65632 (32x2051) -> 2048

Codebook0 Head Out Dim: 2051

Audio Head Shape: torch.Size([31, 1024, 2051])

Backbone max\_seq\_len: 2048

Decoder max\_seq\_len: 2048

Num Codebooks: 32, Audio Vocab Size: 2051

Pre-computed backbone mask: torch.Size([2048, 2048]), device=cuda:0

Pre-computed decoder mask: torch.Size([2048, 2048]), device=cuda:0