

CASE STUDY

Voice-Based Singing Audio Collection for EQ & Compression Algorithm Training: Capturing Linguistic and Musical Diversity



COMPANY

A leading technology company specializing in audio processing and AI-based automation, focused on developing advanced EQ and compression algorithms for the music industry.



INDUSTRY

Music Technology, Audio Processing



USE CASE

Collection of singing audio data for training & evaluating automated EQ & compression systems.



OUR OFFERING

- Data collection & transcription
- Studio-quality recording with real artists across multiple languages & genres

KEY STATS

4 languages:
Chinese, Arabic, Spanish, Russian

10
Singers per language (40 total)

20
Hours of Singing Audio

Audio transcription in **native languages**

Audio format:
48 kHz PCM, mono, WAV

Project Duration
18 Weeks

OVERVIEW

Shaip partnered with a leading technology company to collect diverse singing audio recordings across four prioritized languages: Chinese, Arabic, Spanish, and Russian. The project aimed to provide high-quality data for training AI-based EQ and compression algorithms, which are essential for improving automated audio processing. The collection included 40 participant (10 per language) from various genres, with a focus on studio-quality recordings using diverse microphones & environments.

PROJECT SCOPE

Data Collection

The scope encompassed the collection of singing audio in four targeted languages, recorded by real artists across multiple musical genres. A studio environment was used to ensure high-quality recordings suitable for training AI models.

Key Requirements

- » **Participants:** 10 singers per language, with a balanced gender distribution (50% male, 50% female).
- » **Genres:** A variety of genres, self-identified by the artist, validated for consistency.
- » **Recording Environment:** Studio-quality, with multiple microphone settings (dynamic, condenser).
- » **Audio Format:** 48 kHz PCM, mono, WAV files, with no processing (e.g., no compression, EQ, reverb).
- » **Transcription:** Songs to be transcribed in the language they are sung, with special rules for bilingual songs.
- » **Languages:** Chinese, Arabic, Spanish, Russian

» **Transcription**

- Transcriptions should be provided in the language of the recording (e.g., Hindi lines in Devanagari, followed by English).
- Ensure each segment is no longer than 15 seconds for clarity and accuracy.

» **Audio Recording Requirements**

- Minimum 3 microphone settings per recording session.
- 3 minutes per song, with 3 takes per song, ensuring diverse microphone recordings for each participant.
- Studio-quality acoustic environ

CHALLENGES



Participant Diversity: Ensuring a balanced distribution of singers by gender, voice tone/pitch, and musical genre was a complex challenge.



Data Consistency: Maintaining consistent microphone settings and environment while capturing diverse vocal performances in multiple languages.



Audio Quality Control: Ensuring studio-quality audio without external noise, and accurate transcription in multiple languages.

SOLUTION

Shaip delivered a comprehensive solution to meet the project's requirements by:

- » Recruiting **40 singers** across four languages and ensuring diverse representation in gender, pitch, and musical style.
- » Conducting **studio-quality recordings** with varied microphone types (dynamic, condenser) to capture a wide range of audio data.
- » **Transcribing recordings accurately** in the languages used, following specific rules for bilingual songs.
- » **Consent:** Consent forms will be collected from all participants prior to recording.

THE OUTCOME

The diverse singing audio data collected allowed the client to develop a robust training set for automated EQ and compression algorithms, enhancing the quality of audio processing. The high-quality recordings and detailed metadata ensured that the AI models could handle various musical genres and linguistic complexities. Key Outcomes:

- » **High-quality, diverse audio data** for training AI systems.
- » **Accurate transcription and metadata** for analysis.
- » **A stronger foundation** for AI-based audio processing tools.

DELIVERABLES

- » **20 hours** of studio-quality audio recordings (48 kHz PCM, mono WAV files).
- » **Transcriptions** in the language of the recording.
- » **Metadata:** microphone make/model, DAC/audio interface, singer profile, genre information.
- » **JSON format** for transcription with metadata.

CUSTOMER TESTIMONIAL



Shaip's ability to capture the diversity of musical talent and linguistic richness has been invaluable for the development of our EQ and compression algorithms. Their team ensured that every aspect, from artist recruitment to recording quality, was handled with precision, making this an essential step in refining our automated audio processing systems.

We are truly grateful for the trust and collaboration Shaip has shown throughout the process. Despite our strict and challenging technical requirements, their dedication, hard work, and attention to detail have been outstanding. It has been a pleasure working with a team so committed to delivering excellence

ABOUT THE CLIENT

The client is a leader in the music technology space, specializing in AI-driven solutions for audio processing. Their innovative tools are used by professionals across the music industry to enhance sound quality and automate complex audio tasks.