

# Przeglądarka embeddingów audio - prototyp

Zespół nr 6

Aleksander Gryzik

Jakub Sałek

Rafał Kuśmierz

## Audio Embedding Explorer

 Home

 Pair Analysis

 Model Comparison

 Similarity Ranking

 Pseudo-captioning

 Local Database

## Audio Embedding Explorer

A tool for comparing and visualizing audio-text embeddings, that allows users to upload audio files, enter text, and analyze how embedding models CLAP, OpenL3 and MERT interpret the meaning of the provided data.

Use the sidebar to navigate between views.

Current modules:

- Pair Analysis (CLAP / OpenL3 / MERT)
- Model Comparison on demo data
- Similarity Ranking (Text→Audio / Audio→Text)
- Pseudo-captioning
- Local Database search

## Audio Embedding Explorer

Home

Pair Analysis

Model Comparison

Similarity Ranking

Pseudo-captioning

Local Database

## Pair Analysis

Analyze a single audio-text or audio-audio pair using specified models. This view uses \*\*PCA (Principal Component Analysis)\*\* to visualize embeddings in 2D space — it is the only method for comparing just two samples.

Select embedder model

CLAP

Audio file



Drag and drop file here

Limit 200MB per file • WAV, MP3

Browse files



Lick\_Tenor\_Sax.wav 289.3KB



0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

0:03

Text prompt

White noise static from an untuned TV.

Generate embeddings

Cosine similarity

-0.1841

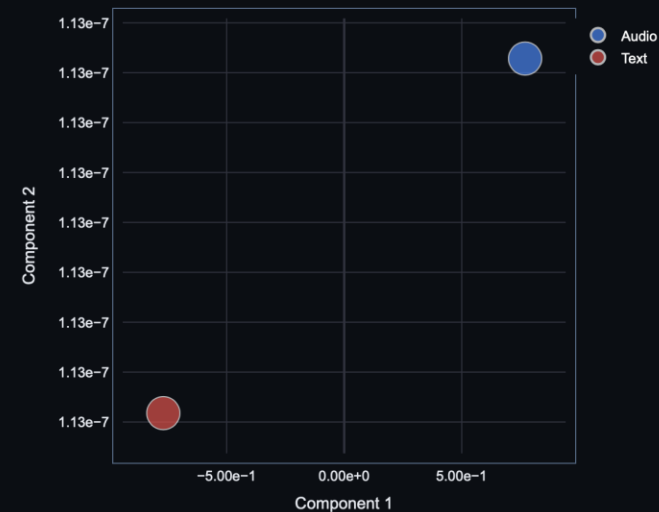
Generate embeddings

Cosine similarity

-0.1841

## 2D PCA projection

2D PCA Projection of Embeddings



## Audio Embedding Explorer

Home

Pair Analysis

Model Comparison

Similarity Ranking

Pseudo-captioning

Local Database

## Pair Analysis

Analyze a single audio-text or audio-audio pair using specified models. This view uses \*\*PCA (Principal Component Analysis)\*\* to visualize embeddings in 2D space — it is the only method for comparing just two samples.

Select embedder model

CLAP

Audio file



Drag and drop file here

Limit 200MB per file • WAV, MP3

Browse files



Lick\_Tenor\_Sax.wav 289.3KB



0:03



Text prompt

Isolated tenor sax phrase, medium register, smooth timbre, studio recording.

Generate embeddings

Cosine similarity

0.6397

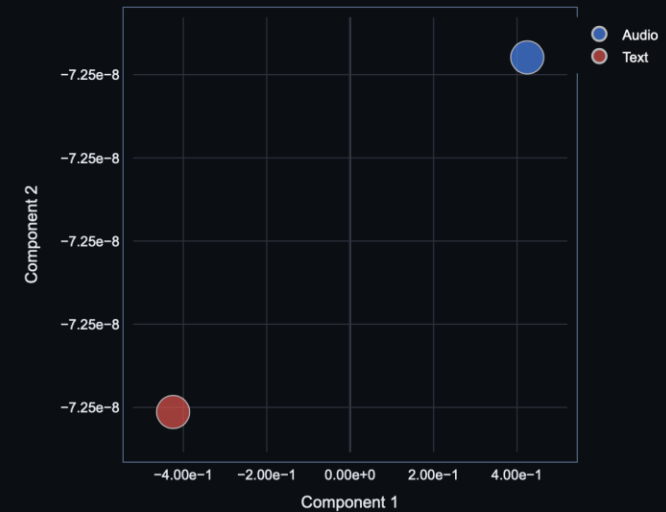
Generate embeddings

Cosine similarity

0.6397

## 2D PCA projection

2D PCA Projection of Embeddings



Audio Embedding Explorer

Home

Pair Analysis

Model Comparison

Similarity Ranking

Pseudo-captioning

Local Database

Pair Analysis

Analyze a single audio-text or audio-audio pair using specified models. This view uses \*\*PCA (Principal Component Analysis)\*\* to visualize embeddings in 2D space — it is the only method for comparing just two samples.

Select embedder model

OpenL3

Audio file 1

Drag and drop file here  
Limit 200MB per file • WAV, MP3

Browse files

bass\_electronic\_018-022-100.wav 128.0KB

0:04

0:04

Speaker icon

Fullscreen icon

Next icon

Audio file 2

Drag and drop file here  
Limit 200MB per file • WAV, MP3

Browse files

bass\_electronic\_018-023-025.wav 128.0KB

0:04

0:04

Speaker icon

Fullscreen icon

Next icon

Generate embeddings

Cosine similarity

0.9920