Sound Dataset Annotation Tool

Daniel Lounsbury

Abstract

The Sound Dataset Annotation Tool is a highly flexible framework for annotating datasets of audio for use in machine learning experiments. It provides an interactive experience that allows users to annotate audio excerpts with semantic tags, such as emotional or timbral descriptors. This tool engages participants and helps researchers to create custom studies, track study progress, and export results. Study designers can configure study parameters, including sounds, descriptors, and annotation flow, to generate high-quality annotations of datasets needed to train machine learning models, such as deep neural networks.

Motivation

Training robust models requires high-quality labeled sound data, but collecting such data is often a slow, difficult, and time-consuming process. We are designing the Sound Dataset Annotation Tool to streamline this process and empower researchers to efficiently generate rich, task-specific datasets for machine learning.

	_id	values	user	updated	style	study
1	67c7770c35e5b9ae0b4b46b0	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,289.0 (1.7T)	best-match	testing
2	67c7770c35e5b9ae0b4b46b1	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,292.0 (1.7T)	best-match	testing
3	67c7770c35e5b9ae0b4b46b2	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,295.0 (1.7T)	best-match	testing
4	67c7770c35e5b9ae0b4b46b3	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,299.0 (1.7T)	best-match	testing
5	67c7770c35e5b9ae0b4b46b4	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,302.0 (1.7T)	best-match	testing
6	67c7770c35e5b9ae0b4b46b5	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,304.0 (1.7T)	best-match	testing
7	67c7770c35e5b9ae0b4b46b6	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,307.0 (1.7T)	best-match	testing
8	67c7770c35e5b9ae0b4b46b7	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,310.0 (1.7T)	best-match	testing
9	67c7770c35e5b9ae0b4b46b8	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,314.0 (1.7T)	best-match	testing
10	67c7770c35e5b9ae0b4b46b9	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,317.0 (1.7T)	best-match	testing

Fig. 1: Illustration of annotations stored in the database.

Oregon State University

Users / Study Annotation

Users are tasked to annotate sound samples.

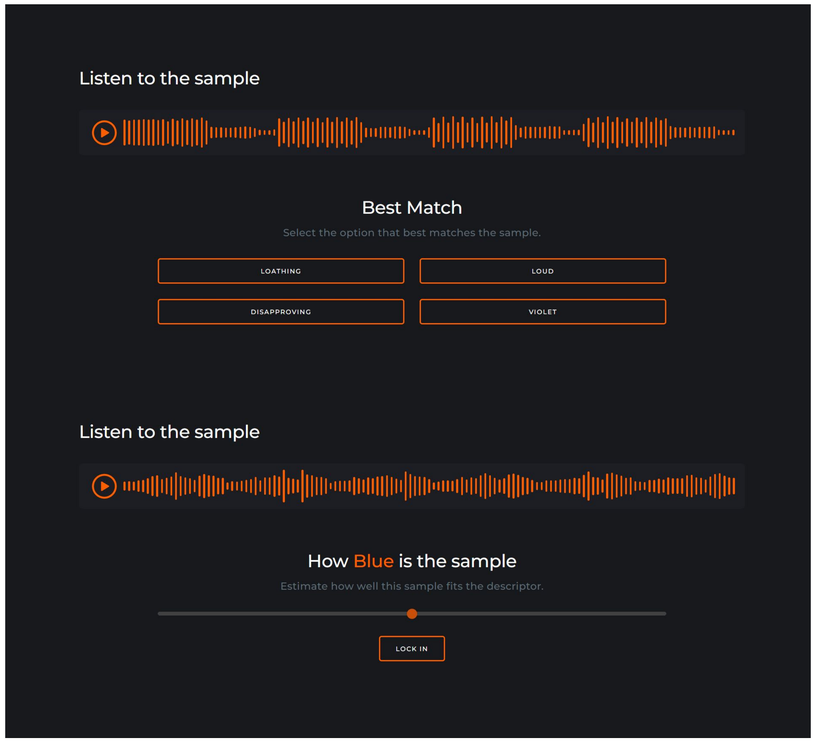

Fig. 2: Examples of sound annotation pages presented to users.

Study Modes

The system supports six different annotation styles:

- **Best Match**:
  Pick one descriptor that best fits the sound.
- **Pick All**:
  Select all the descriptors that apply to the sound.
- **Compare**:
  Choose which sound best fits the descriptor.
- **Strength**:
  Rate how well a descriptor fits the sound.
- **Opposites**:
  Pick the best fit from a contrasting pair.
- **Forced Match**:
  Match each sound to one descriptor, no repeats.

Admins/Researchers

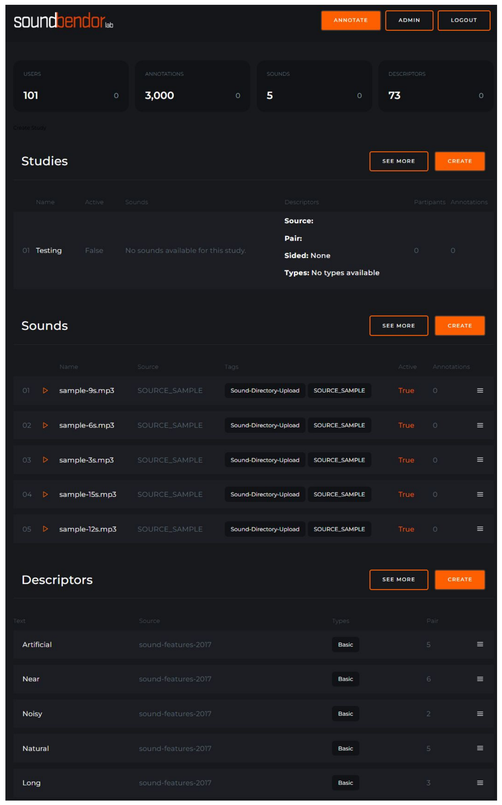

Fig. 3: Screenshot of the admin management page.

Administrators can:

- Create new and custom studies.
- Add or create custom descriptors.
- Upload sounds and necessary metadata.
- Export annotation data from studies.

Future Work

- Add gamification and user incentives.
- Build tools for analytics of collected data.
- Design a large scale study to validate tool.

soundbendor lab