

Oregon State University

Sound Dataset Annotation Tool

Daniel Lounsbury

Project Overview

Oregon State University

- Web-based system for collecting music annotations
- Researchers configure studies with custom descriptors and sound samples
- Users provide subjective audio classifications
- Data collection for machine learning model training

soundbendor lab

Research Value Proposition

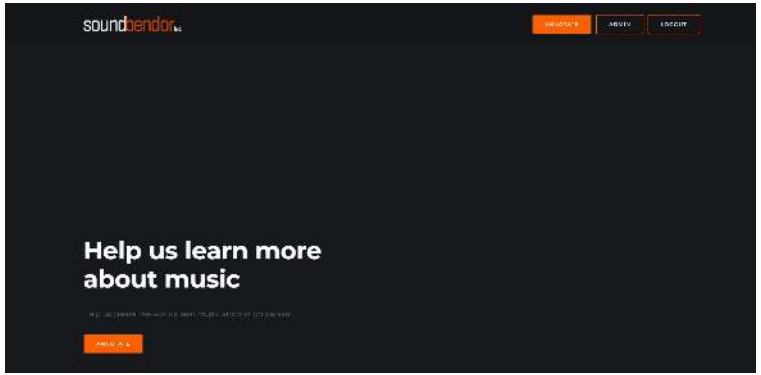
Oregon State University

- Addressing Audio Classification Challenges
- Subjective audio descriptors require human perception
- Crowdsourcing scales data collection efficiently
- Gamification increases participant engagement
- Standardized data format enables research reproducibility
- Multi-study platform supports diverse research needs

soundbendor lab

System Architecture

- Frontend: React web application
- Backend: Node.js REST API
- Database: MongoDB for data persistence
- Deployment: AWS cloud infrastructure
- Authentication: Session-based user management



SDAT Main Landing Page

Data Management Strategy

Oregon State University

- Studies: Configuration and metadata
- Descriptors: Annotation categories and parameters
- Audio Files: Sample storage and streaming
- User Data: Accounts and session management
- Annotations: User responses and timestamps



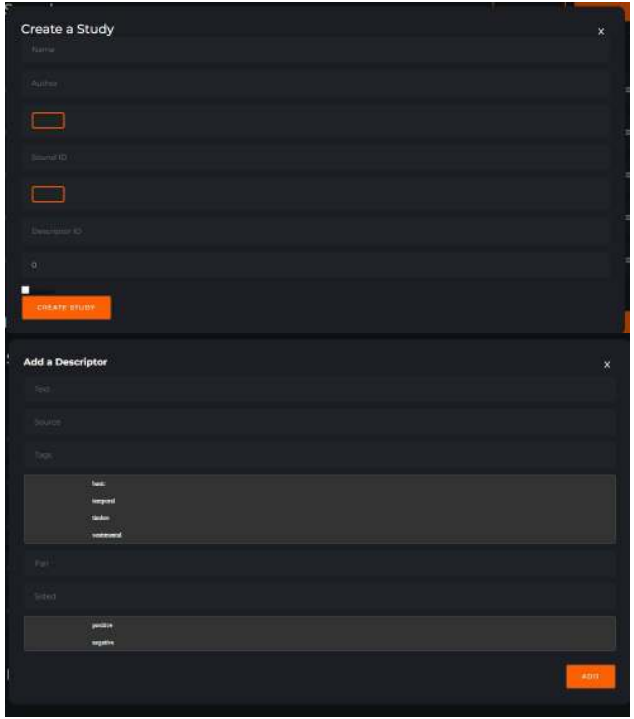
	_id	values	user	updated	style	study
1	67c7770c35e5b9ae0b4b46b0	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,289.0 (1.7T)	best-match	testing
2	67c7770c35e5b9ae0b4b46b1	Array[1]	3a1f4004-bda5-431b-9cbe-313c3b5b6853	1,741,125,388,292.0 (1.7T)	best-match	testing
3	67c7770c35e5b9ae0b4b46b2	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

Illustration of annotations stored in the database.

soundbendor lab

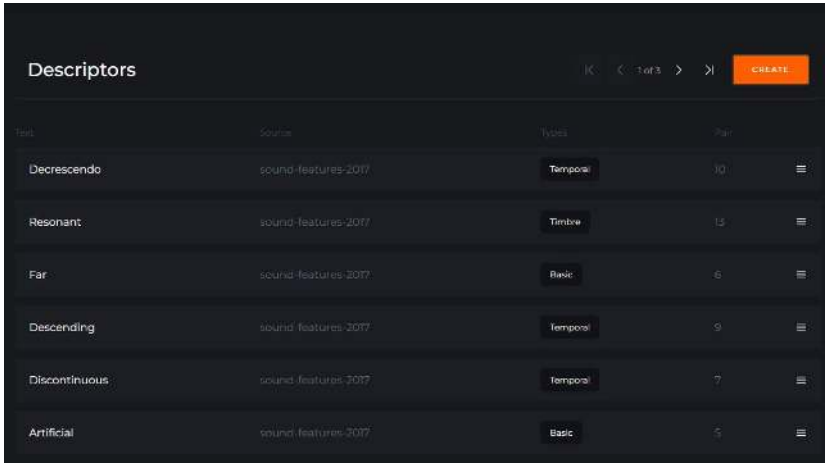
Current Progress

- Admin dashboard with data aggregation
- Study creation dialog implementation
- Descriptor creation interface
- Admin list pages with pagination (Studies, Descriptors, Sounds)
- User logout functionality
- AWS production deployment
- Admin-only access enforcement
- Screenshots: Admin dashboard, creation dialogs, list interfaces



Development Roadmap

- Next Phase Priorities
- Individual resource detail pages implementation
- List filtering and search functionality
- Complete CRUD operations (edit/delete)
- Enhanced CSV export capabilities
- Advanced study configuration options
- User experience optimization



No CSV export on for descriptors

soundbendor lab



Oregon State University
Cascades