

[Canva Link here!](#)

EVERETT-ALAN HOOD

SURROUND SOUND



Two eyes, two ears.

Computer Vision is one of the most powerful tools of modern technology. The ability to process imagery enables developers to push towards the future and enforce safety.

Our ears do the same, so where's the drive towards auditory awareness?



The State of Auditory AI

There's a focus on human speech and music entertainment.

Besides medical research, auditory AI is not being promoted in other practical fields.

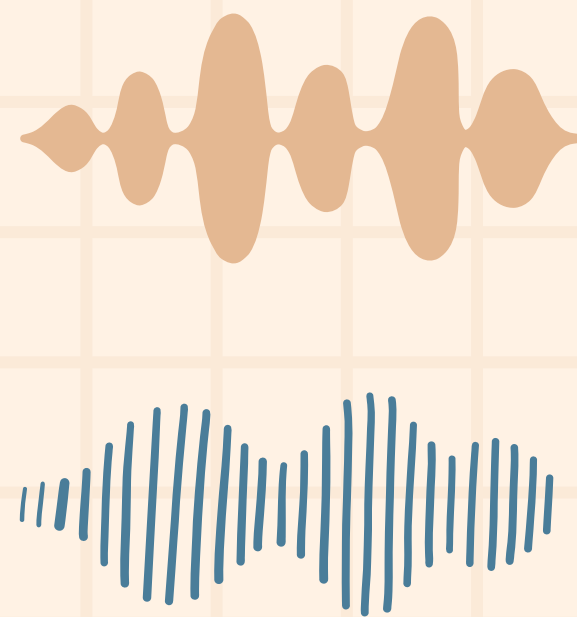
“Existing efforts of audio-text alignment in LLMs mainly focus on content information—refers to the explicit meaning and linguistic structure.” -arXiv 2025

“Generative AI for music creation will gradually increase as more DMSPs explore AI driven production and collaboration. Enhanced Personalization will gradually grow in sophistication as AI constantly refines user experience.” -Frontiers 2025

Much of the research among the leading STEM institutions relates to AI in audiology including Harvard, MIT, Columbia, and the University of Washington.

Use Cases & Applications

- Accessibility & Assistance
- Social Awareness
- Noise Pollution Detection
- Autonomous Robotics
- Track Behavioral Changes
- Emergency Response



ML 581 Objective

- Create a model that classifies a user's environment from real-time audio.
- Predict likely outcomes of the current scenario based on surrounding context.

Rough Idea

- Clean raw audio & metadata of the human experience
- Use embeddings to train multi-dimensional classification system
- Record at least five seconds of the user's microphone
- Predict the current scenario

Datasets

- industrial and natural ambience
- indoor ambience
- languages & emotional tone
- animal sounds
- music & activities
- foley samples

HuggingFace, Zenodo, Kaggle

Potential Attributes

- WAV/FLAC Format
- Stereo (if possible)
- 16kHz
- Description
- Translation
- Timestamp
- Tags
- LUFS & DB



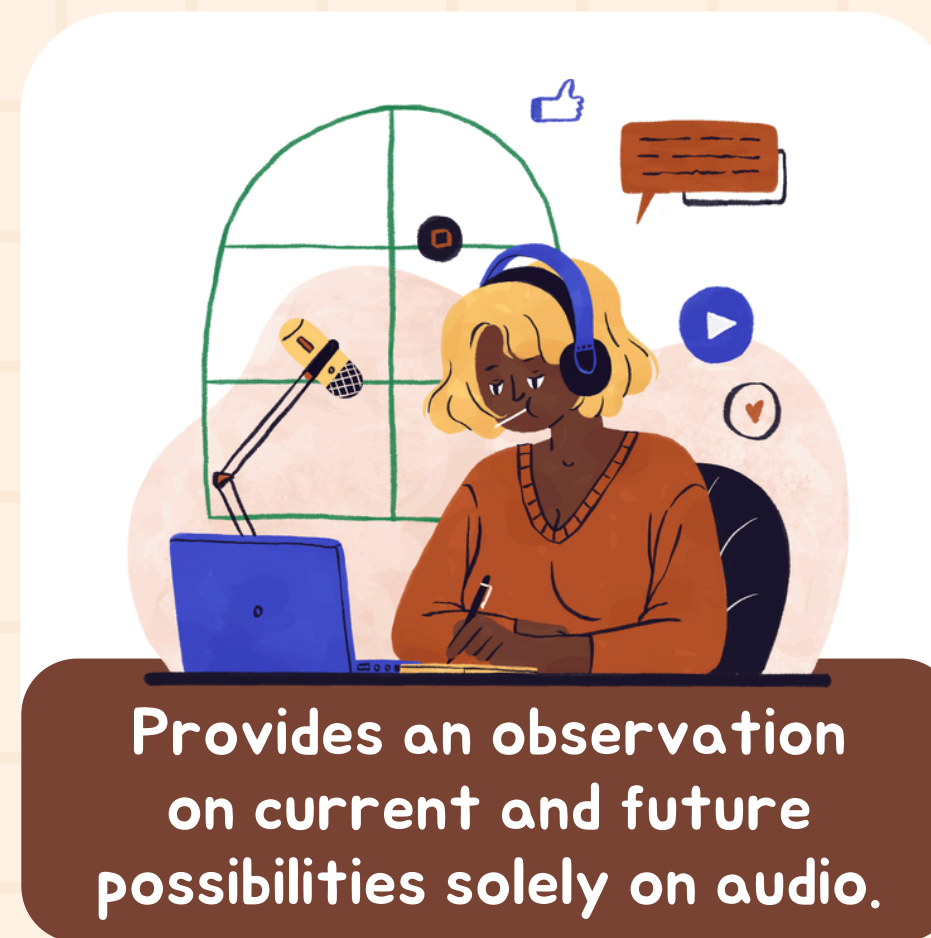
Sample Example

```
{ "filename": "urban_street_01.wav",  
  "format": "WAV",  
  "sample_rate": 16000,  
  "channels": 2,  
  "tags": ["car_horn", "footsteps",  
  "engine"],  
  "lufs": -18.3,  
  "timestamp": "2025-10-06T14:23:10",  
  "desc": "Busy street with passing  
cars and footsteps." }
```





Audio is info. Information is revelation.



**We want our programs to cooperate with us and keep us safe.
Why not put technology in our shoes?**

THANK YOU!