SPECTRA:

Personalizable Sound Recognition for DHH Users through Interactive Machine Learning

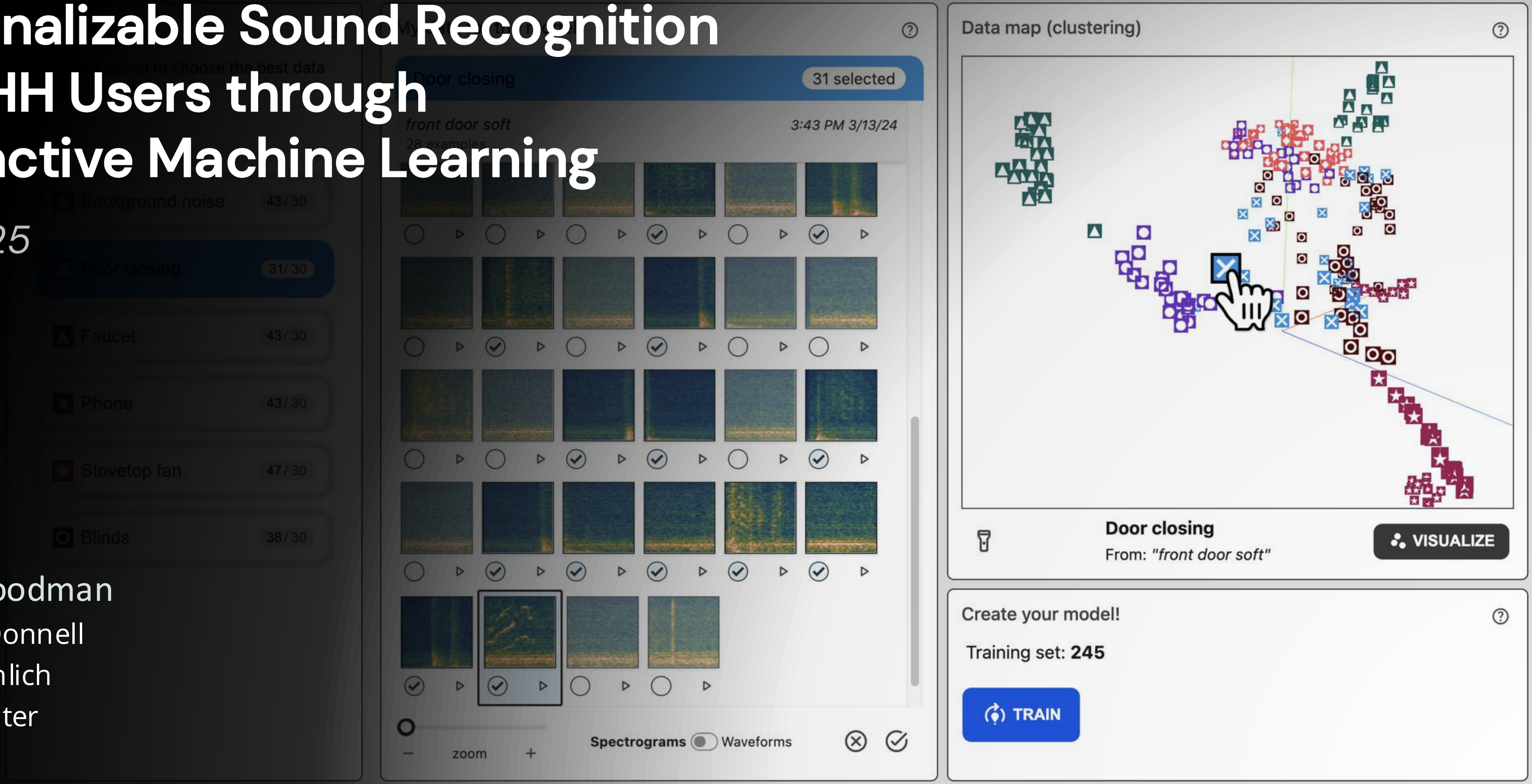
CHI 2025

UNIVERSITY of WASHINGTON

HUMAN CENTERED DESIGN & ENGINEERING

inclusive design lab

DUB



Steven Goodman

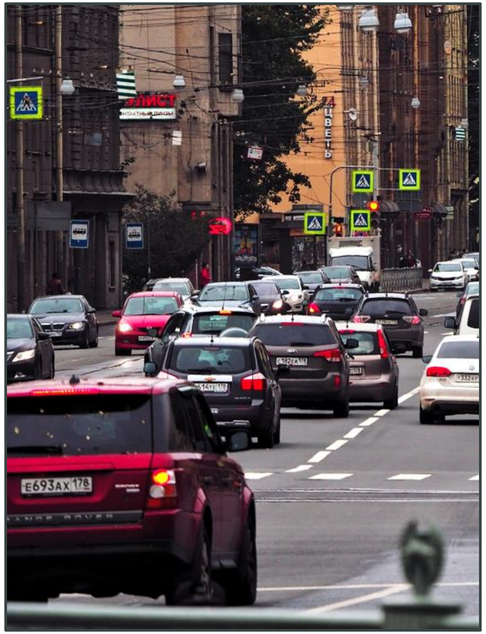
Emma McDonnell

Jon E. Froehlich

Leah Findlater

Introduction

Sound carries rich information about the world around us



Introduction

Sound carries rich information about the world around us

But it may not be accessible to people who are Deaf and hard of hearing (DHH)

Current tools are inadequate

Sound recognition tools have proliferated in the research literature.

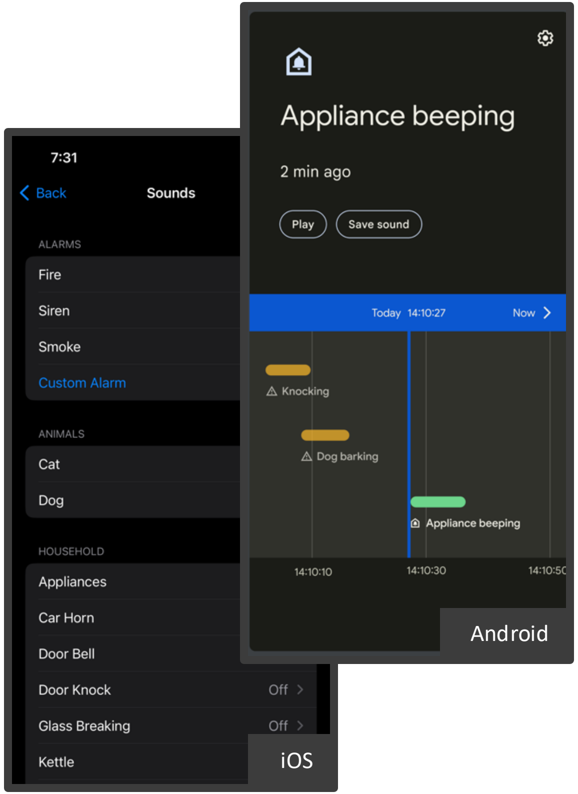
E.g., Bragg et al. (ASSETS 2016); Sicong et al. (IMWUT 2017); Jain et al. (ASSETS 2020)

Android & iOS offer automatic sound recognition using pre-trained models.

- ~15 sound categories: appliances, alarms, pets

Surveyed DHH users expressed dissatisfaction with accuracy & available sounds.

Jain et al. (CHI 2022)

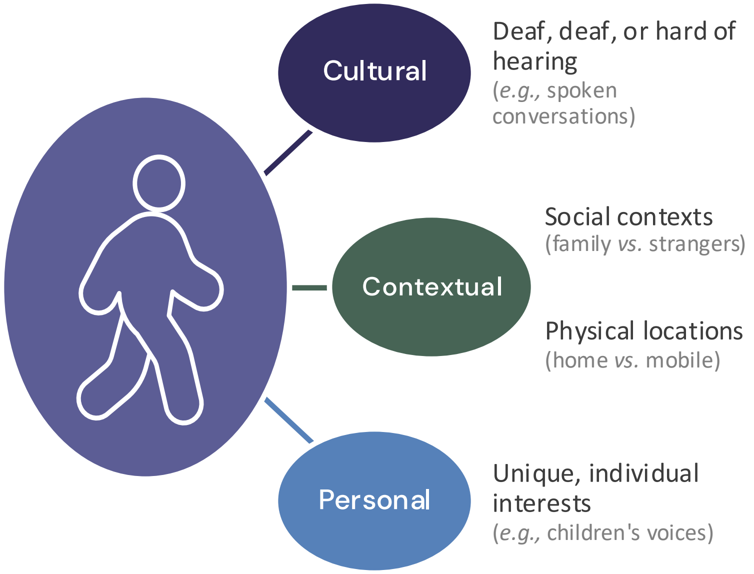


DHH users have diverse needs

Different factors influence sound preferences among DHH individuals.

A "one-size-fits-all" sound awareness solution may not be tenable.

Bragg et al. (ASSETS 2016); Findlater et al. (CHI 2019); Jain et al. (CHI 2019); Matthews et al. (BIT 2006)



In need of personalized support

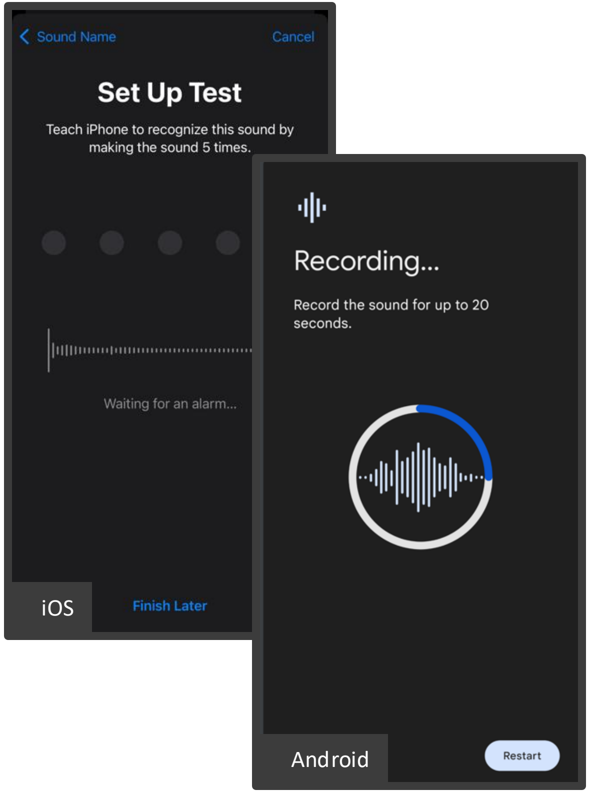
Personalization can better address the varied needs of DHH users.

Currently, users can filter alerts and extend the pre-trained model with their own recordings.

- iOS: fine-tuning existing categories
- Android: adding custom categories

This automatic approach is fast and easy but lacks transparency and control—which may limit trust and long-term use.

Drozdal et al. (IUI 2020)



Interactive ML as a solution?

Interactive ML is promising for accessibility applications, but assumes end-users have domain expertise.

Kacorri et al. (CHI 2017); Sosa-Garcia & Odone (TACCESS 2017)

Gaps remain within IML for sound recognition:

- Nakao *et al.* - lacked visualizations for understanding data.
- Goodman *et al.* - only explored data collection.

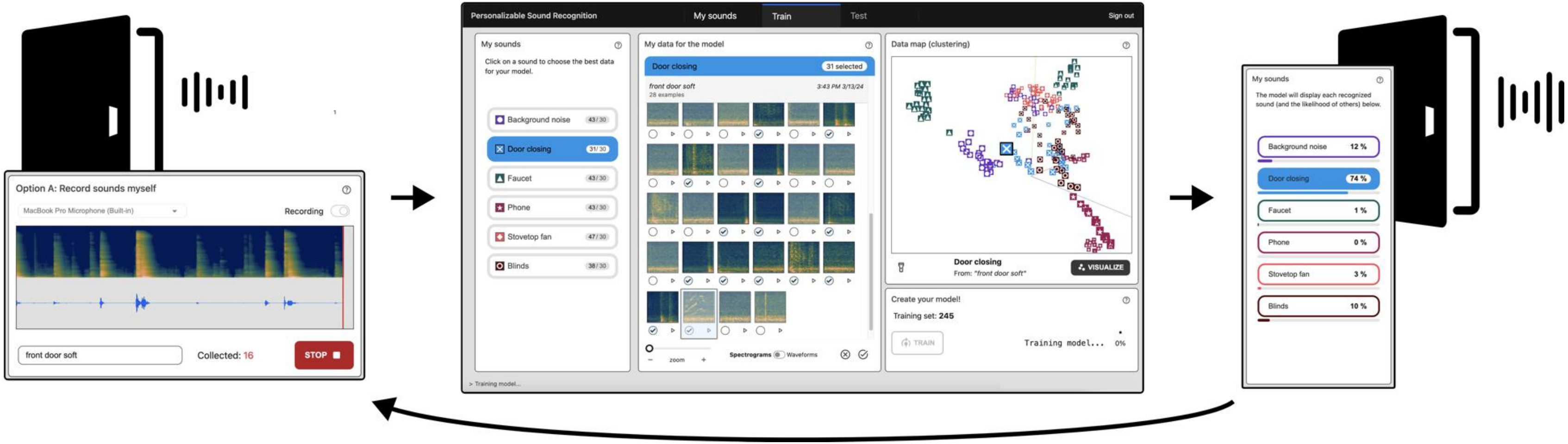
Goals of our research:

- Investigate IML's impact on DHH users' perspectives.
- Identify effective mechanisms for personalizing these systems.

Nakao et al. (NordiCHI 2020)

Goodman et al. (IMWUT 2021)

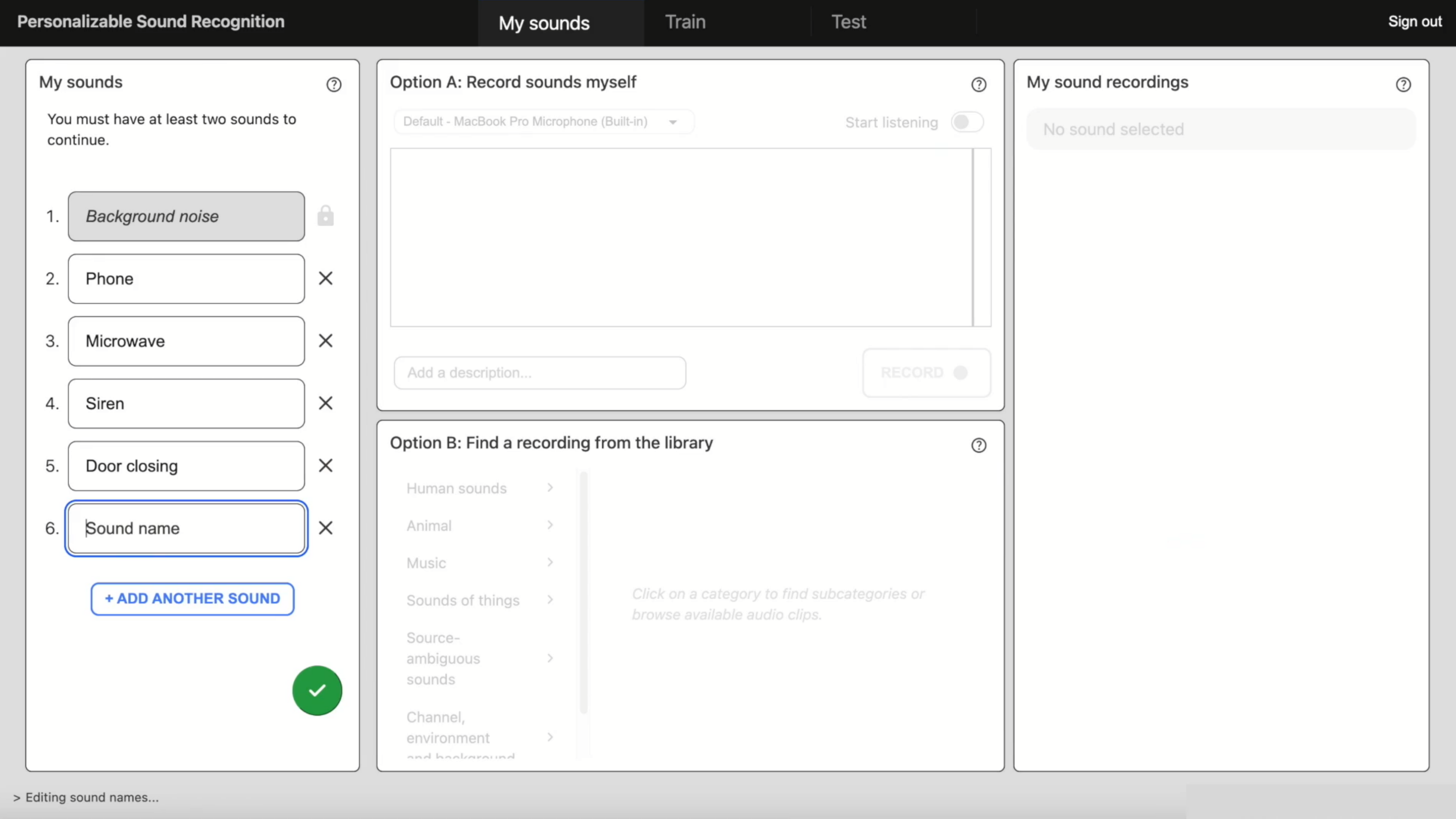
Overview of the SPECTRA pipeline



1. Planning & Data collection

2. Data curation & Model training

3. Testing & Assessment


Personalizable Sound Recognition
My sounds
Train
Test
Sign out
My sounds
You must have at least two sounds to continue.
1. Background noise
2. Phone
3. Microwave
4. Siren
5. Door closing
6. Sound name
+ ADD ANOTHER SOUND
Option A: Record sounds myself
Default - MacBook Pro Microphone (Built-in)
Start listening
Add a description...
RECORD
Option B: Find a recording from the library
Human sounds
Animal
Music
Sounds of things
Source-ambiguous sounds
Channel, environment
Click on a category to find subcategories or browse available audio clips.
My sound recordings
No sound selected
> Editing sound names...

Research Questions

1. How do DHH users engage with SPECTRA to train a personalized sound recognition model?

2. How does interacting with SPECTRA affect DHH users' perspectives on sound models and confidence with custom training?

User evaluation (120 min)

12 DHH participants

- Any tech. experience level
- Moderate confidence in ML concepts
  (7-point scale, avg.=4.8, range=3-6)
- Five w/ hands-on experience

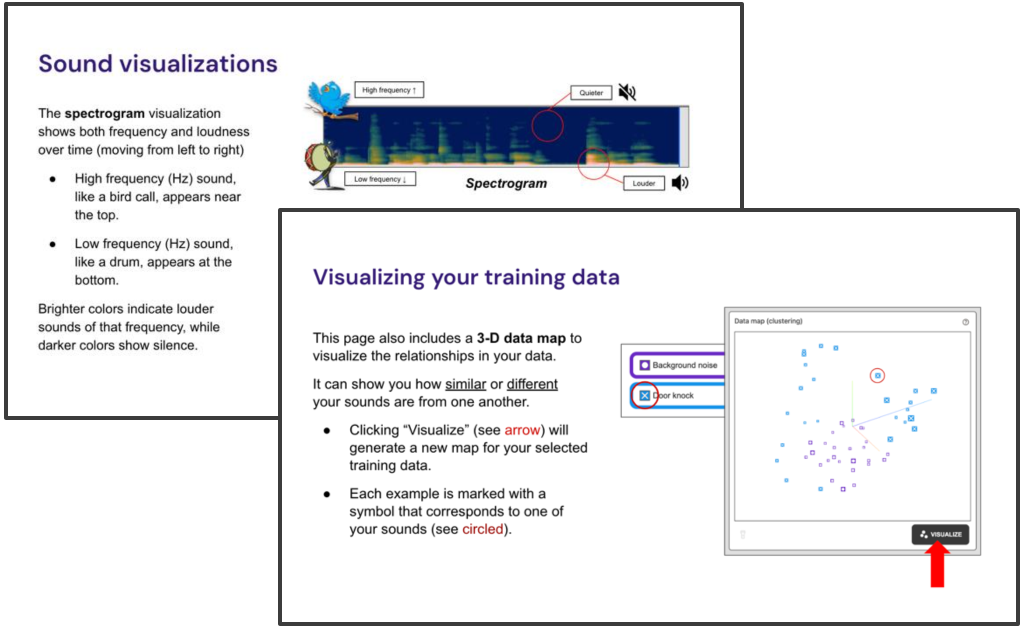
User evaluation (120 min)

12 DHH participants

- Any tech. experience level
- Moderate confidence in ML concepts
  (7-point scale, avg.=4.8, range=3-6)
- Five w/ hands-on experience

Tutorial & Interview
(30 min)

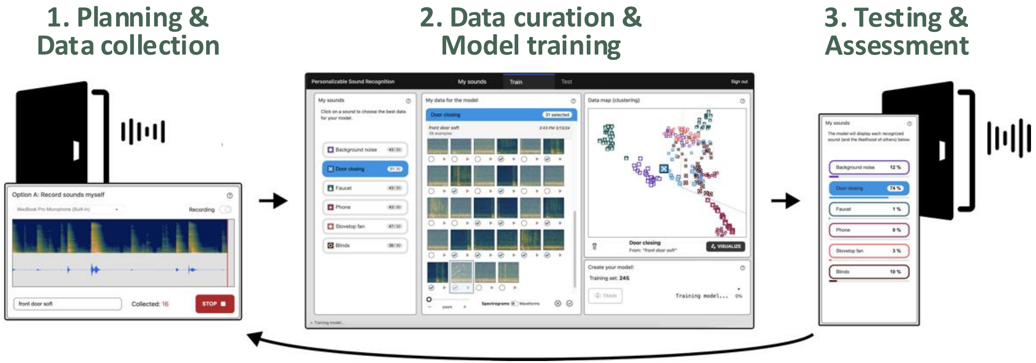
- Introduce SPECTRA and sound recognition concepts
- Capture pre-use expectations



User evaluation (120 min)

12 DHH participants

- Any tech. experience level
- Moderate confidence in ML concepts
  (7-point scale, avg.=4.8, range=3-6)
- Five w/ hands-on experience



Tutorial & Interview
(30 min)
- Introduce SPECTRA and sound recognition concepts
- Capture pre-use expectations

System Use
(60 min)
- Train model for six sounds
- Collect data, build training dataset, test

User evaluation (120 min)

12 DHH participants

- Any tech. experience level
- Moderate confidence in ML concepts
  (7-point scale, avg.=4.8, range=3-6)
- Five w/ hands-on experience

Tutorial & Interview (30 min)
- Introduce SPECTRA and sound recognition concepts
- Capture pre-use expectations

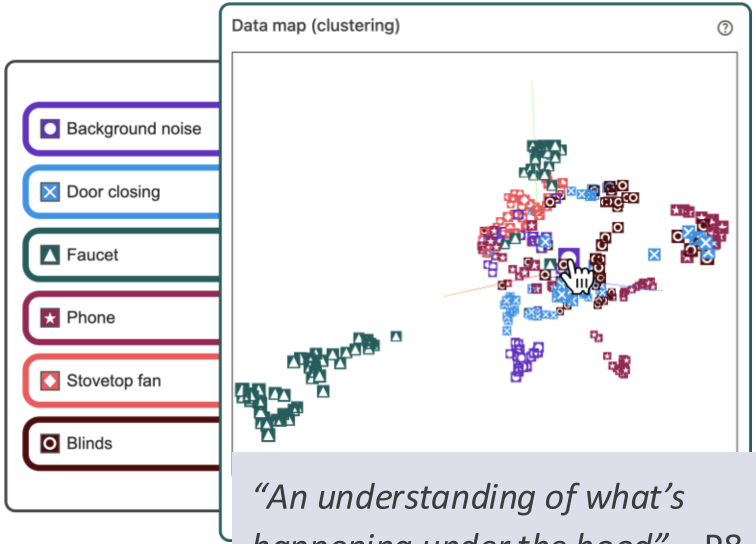
System Use (60 min)
- Train model for six sounds
- Collect data, build training dataset, test

Semi-structured Interview (30 min)
- Reflect on the experience
- Capture post-use perspectives

Findings

#1: Insights through clustering

P10's unfiltered data



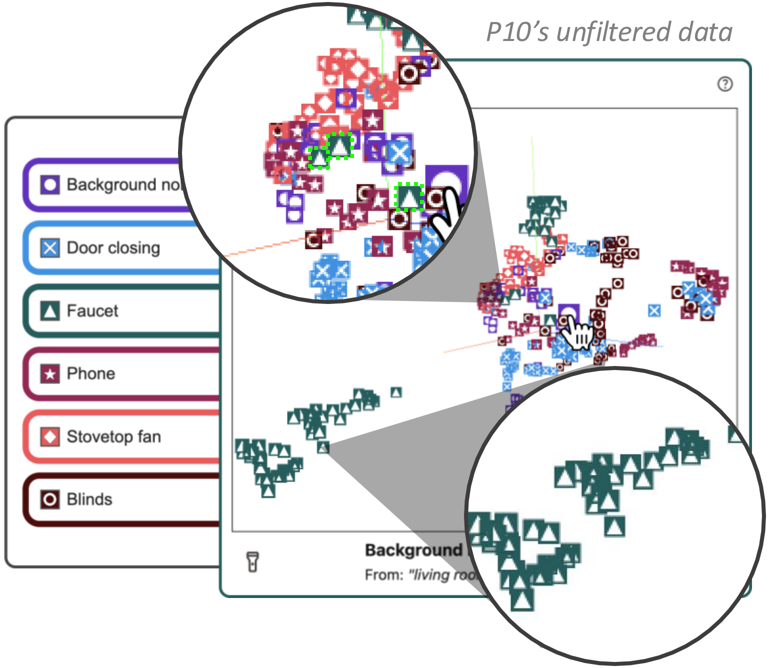
"An understanding of what's happening under the hood" – P8

Clustering was deemed critical to the IML process to:

- Understand similarities among sounds
- Troubleshoot sources of misclassification (via overlap)
- Highlight the most distinct sounds
- Identify outliers and iteratively refine training dataset

Findings

#1: Insights through clustering

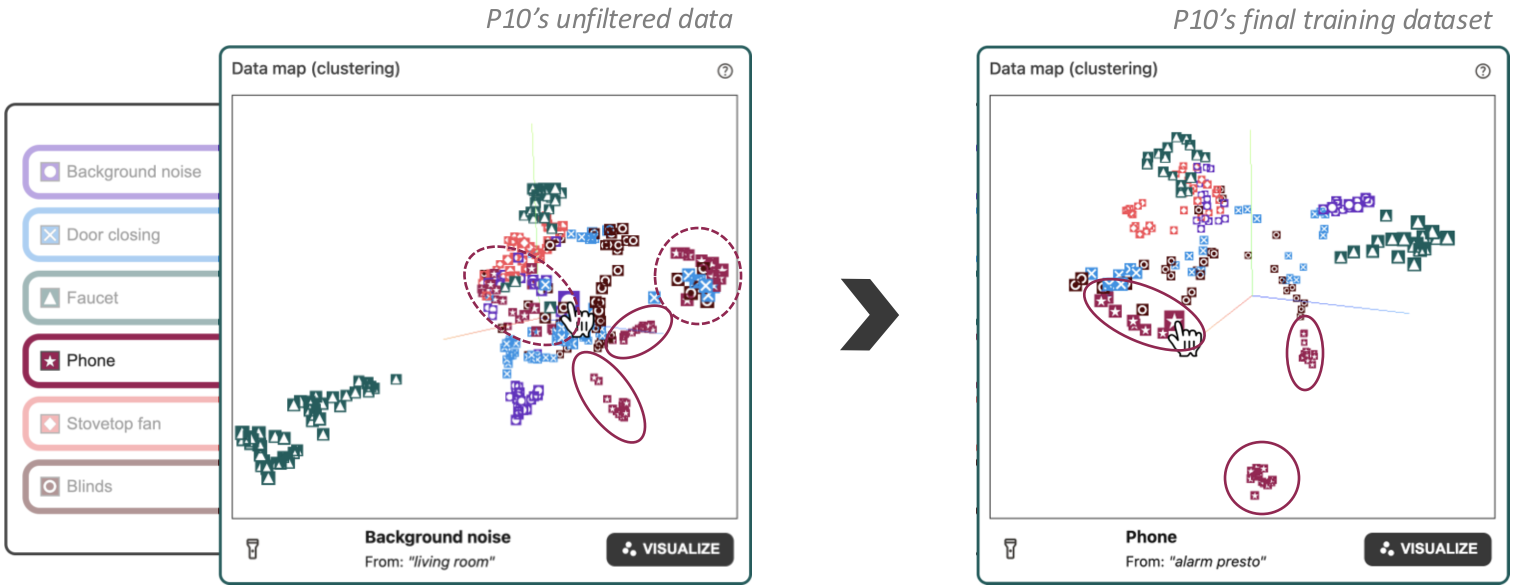


Clustering was deemed critical to the IML process to:

- Understand similarities among sounds
- Troubleshoot sources of misclassification (via overlap)
- Highlight the most distinct sounds
- Identify outliers and iteratively refine training dataset

Findings

#1: Insights through clustering



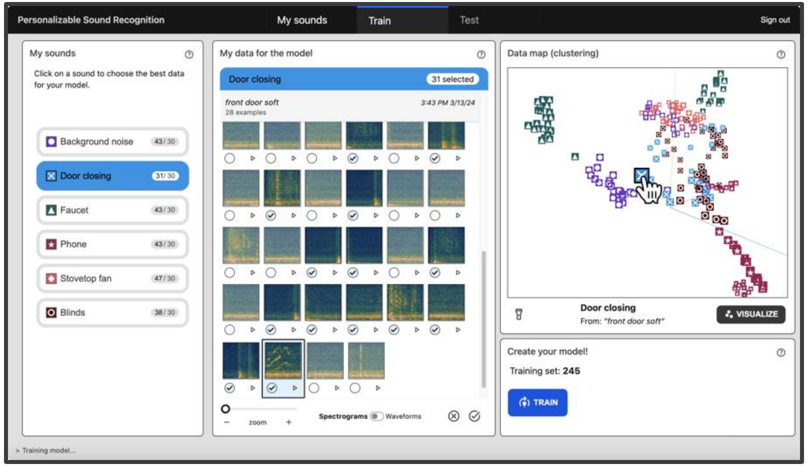
"[It] was satisfying to see, 'Okay, like it's actually working; what I'm doing.'" - P1

Findings

#2: Interplay of Visualizations

Participants combined multiple information streams to make decisions about their models.

- Clustering: High-level view of data structure & relationships
- Waveform: Intuitive, glanceable, good for quick assessment
- Spectrograms: Less intuitive, but useful for in-depth analysis (for some)
- Annotations: Provide context, aid in recall, support deeper understanding

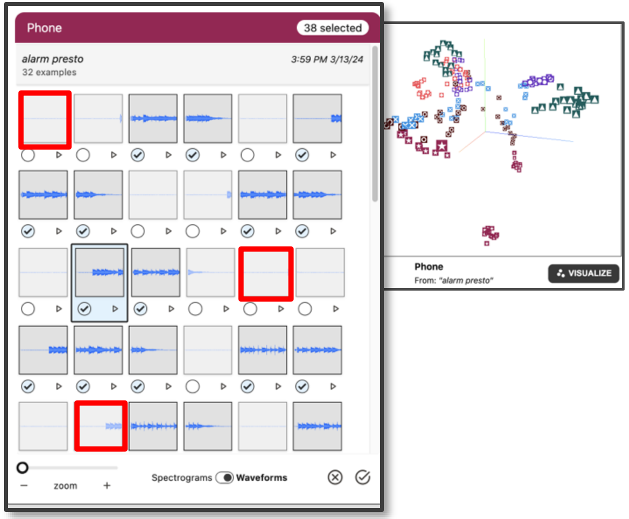


Findings

#2: Interplay of visualizations

Training strategy A *(Example-centric)*

Analysis via example icon, clustering for monitoring



P4 on the waveform’s glanceability:

“The background noise [vs.] whenever I was talking,

Being able to figure out which [example] was which—I think that was really helpful.”

Findings

#2: Interplay of visualizations

Training strategy A *(Example-centric)*

Holistic analysis via example icon, clustering for monitoring

"I was driven by what I was seeing in the chart [...] to eliminate some edge cases and anomalies.

Everything is [shown] together.

In [the selection panel], I have to compare one by one" – P11

Training strategy B *(Clustering-centric)*

Clustering as interactive flagging tool, example icon for targeted analysis

#3: Balancing engagement & efficiency

Responses indicated interactive ML promoted understanding and confidence, but the process was time-consuming.

All participants trained just one model due to time limits or fatigue.

- Workflow requires too many interactions to produce a useful result

> *“The unchecking [was] not my favorite; […] It just ate up time.”* – P3

Findings

#3: Balancing engagement & efficiency

Participants shared their own ideas to improving the process:

- Adding custom classes to an existing model *(e.g.*, Android)
- Automation (*e.g.,* background noise removal)
- *In situ* help (*e.g.,* P7: *"text reminders"* suggesting problematic examples)

How to optimize training process while continuing to support meaningful engagement with the model remains an open question.

UNIVERSITY of WASHINGTON

HUMAN CENTERED DESIGN & ENGINEERING

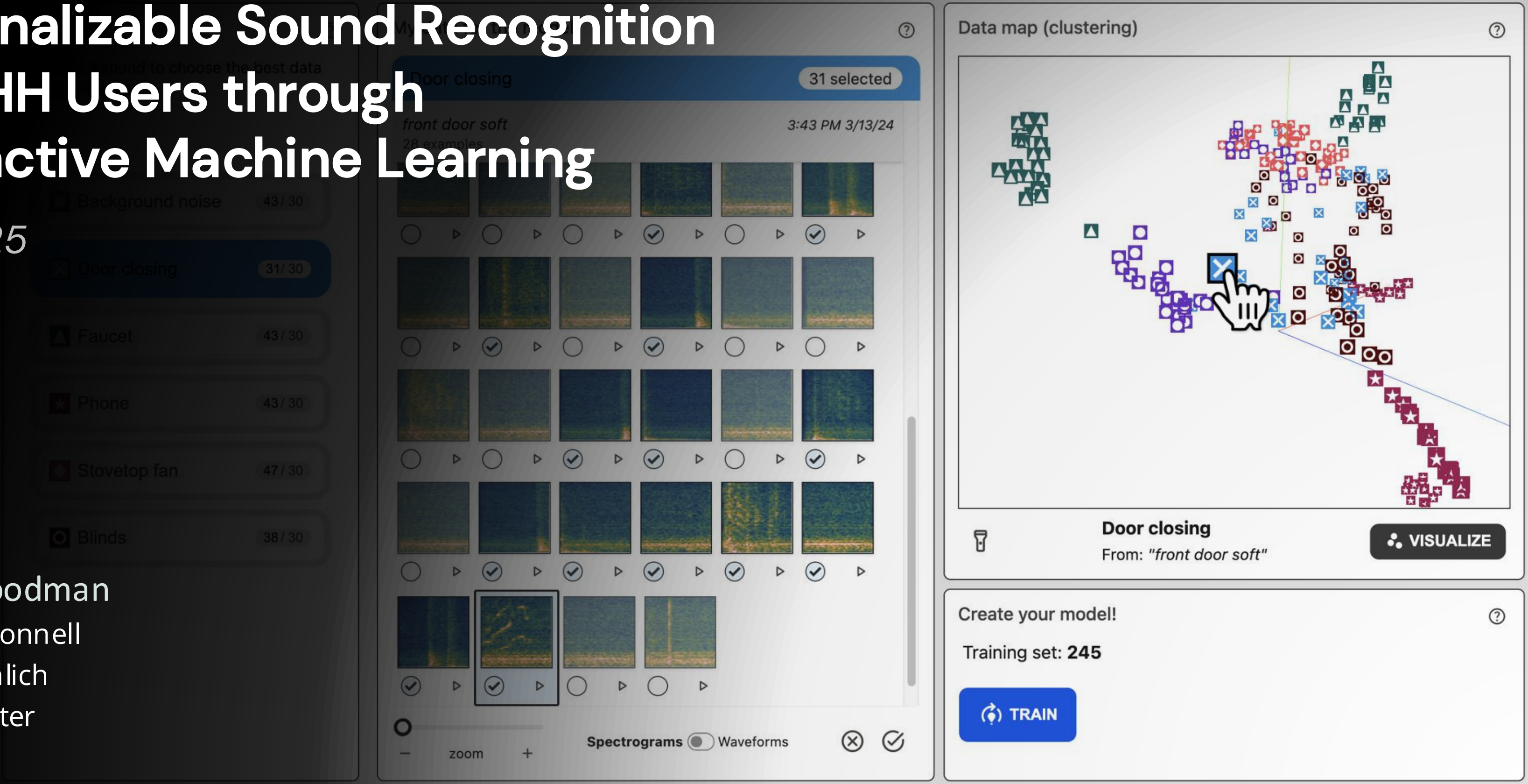
inclusive design lab

DUB

SPECTRA:

Personalizable Sound Recognition for DHH Users through Interactive Machine Learning

CHI 2025



Steven Goodman
Emma McDonnell
Jon E. Froehlich
Leah Findlater

Questions?