Freesound Datasets:
A platform for the creation of open audio datasets

Eduardo Fonseca, Jordi Pons, Xavier Favory, Frederic Font, Dmitry Bogdanov, Andres Ferraro, Sergio Oramas, Alastair Porter and Xavier Serra

name.surname@upf.edu

Music Technology Group, UPF, Barcelona, Spain

audio commons

EXCELENCIA MARÍA DE MAEZTU

upf. Universitat Pompeu Fabra Barcelona

MTG Music Technology Group

Introduction

◆ Datasets are key factor in **sound and music computing**. However:
  - lack of **transparency** in their creation
  - not completely **open/shareable**
  - unclear mechanisms for **error ammendment**
  - **not large enough** for current machine learning needs
◆ Freesound Datasets: online platform for the collaborative creation of open audio datasets

Our vision

◆ **Transparency**: workflow in creation process & content exploration
◆ **Openness**: free distribution & reuse
◆ **Dynamism**: criticism & proposals in the platform
◆ **Sustainability**: community as continuous source of information
◆ **How?**

freesound

The Freesound Datasets platform - https://datasets.freesound.org/

We plan to host several datasets supporting:
◆ **Annotation**:
  - ground truth generation
  - volunteering-based crowdsourcing
  - guidelines & training phase
  - quality control
◆ **Dataset exploration**:
  - browsing sounds by categories
  - interactive visualizations
◆ **Discussion**:
  - platform: improvements & functionalities
  - datasets: content, collection process, taxonomy
◆ **Dataset releases:**
  - time-evolving (quality & quantity)
  - dataset versioning
◆ **Download:**
  - waveforms, features, ground truth, metadata
  - Creative Commons licenses
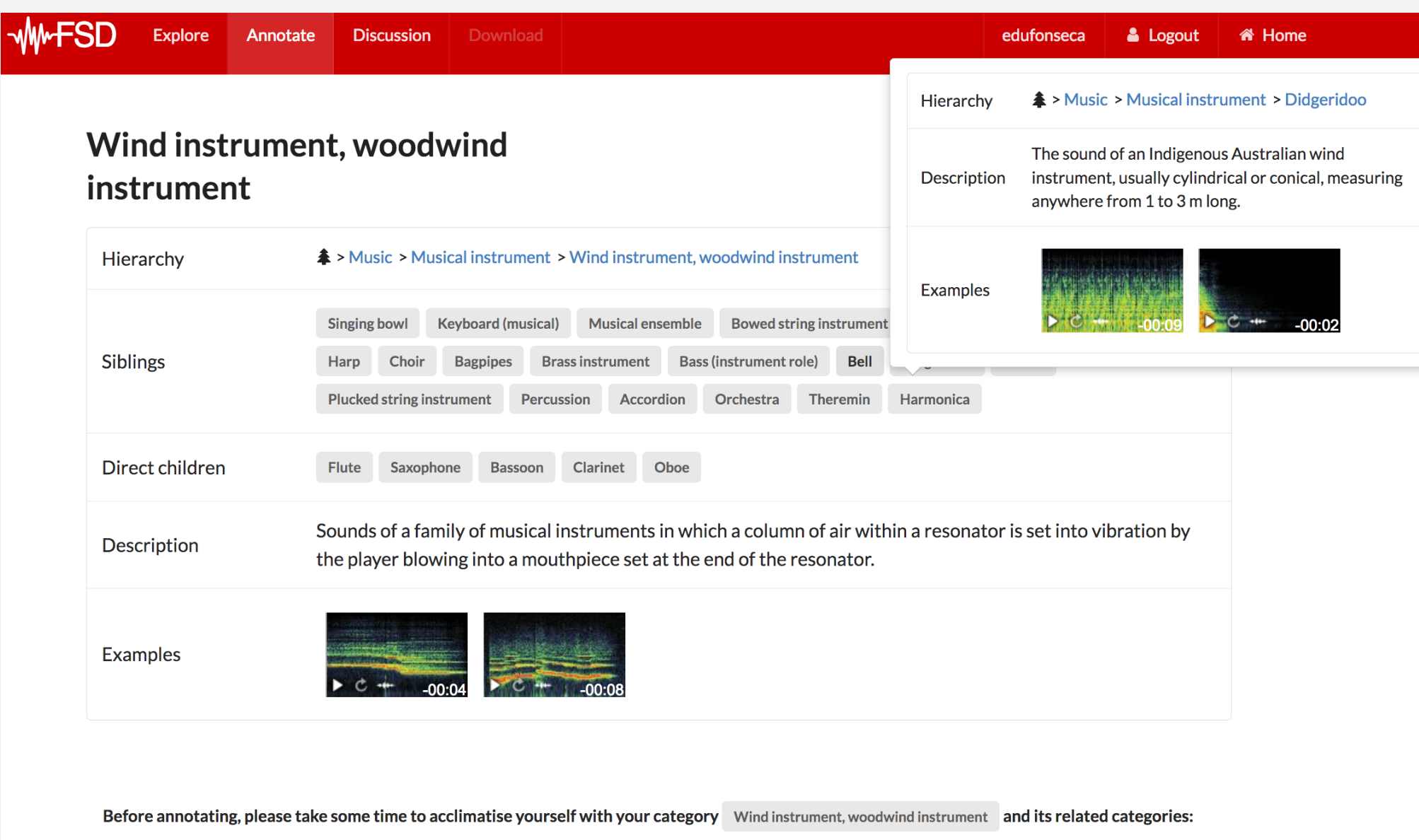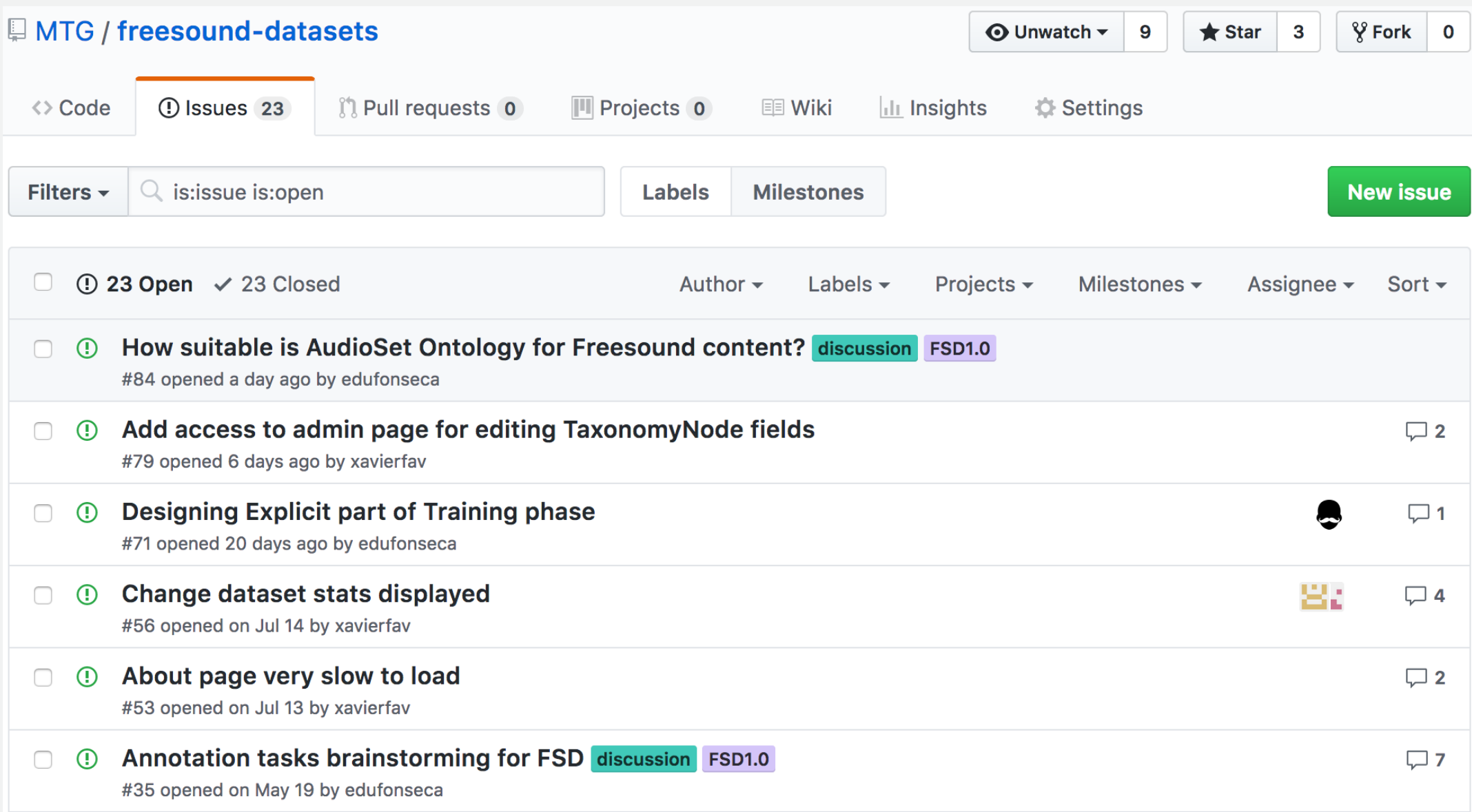
Training phase before annotation
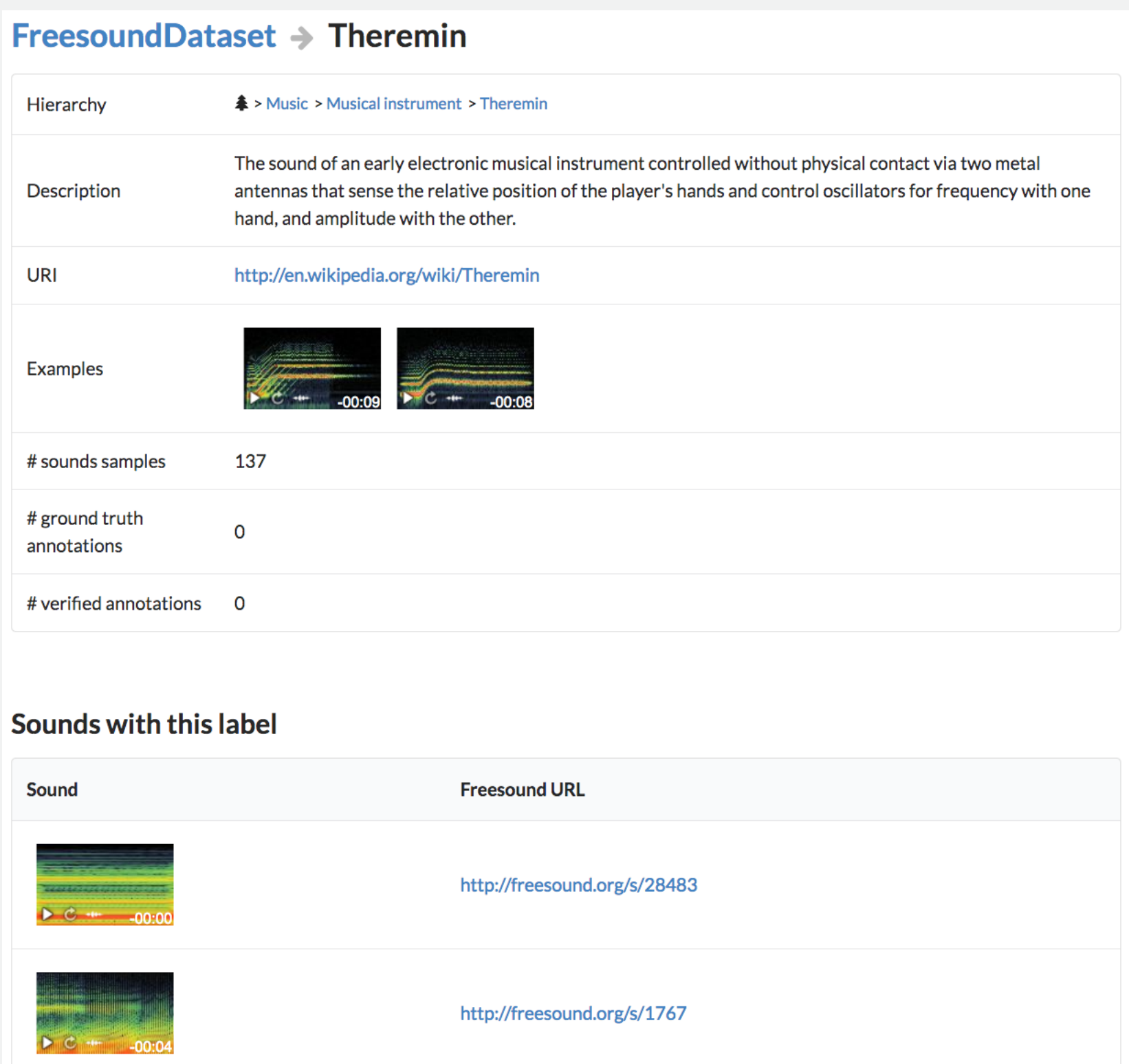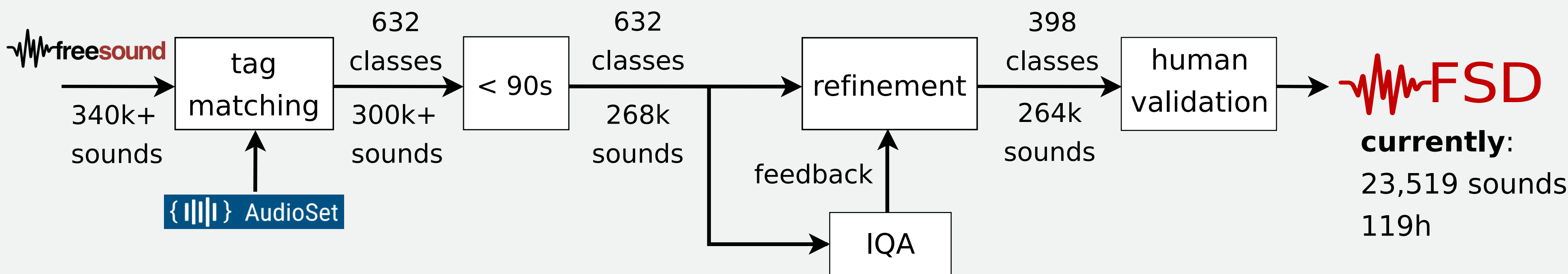
Exploration of a category
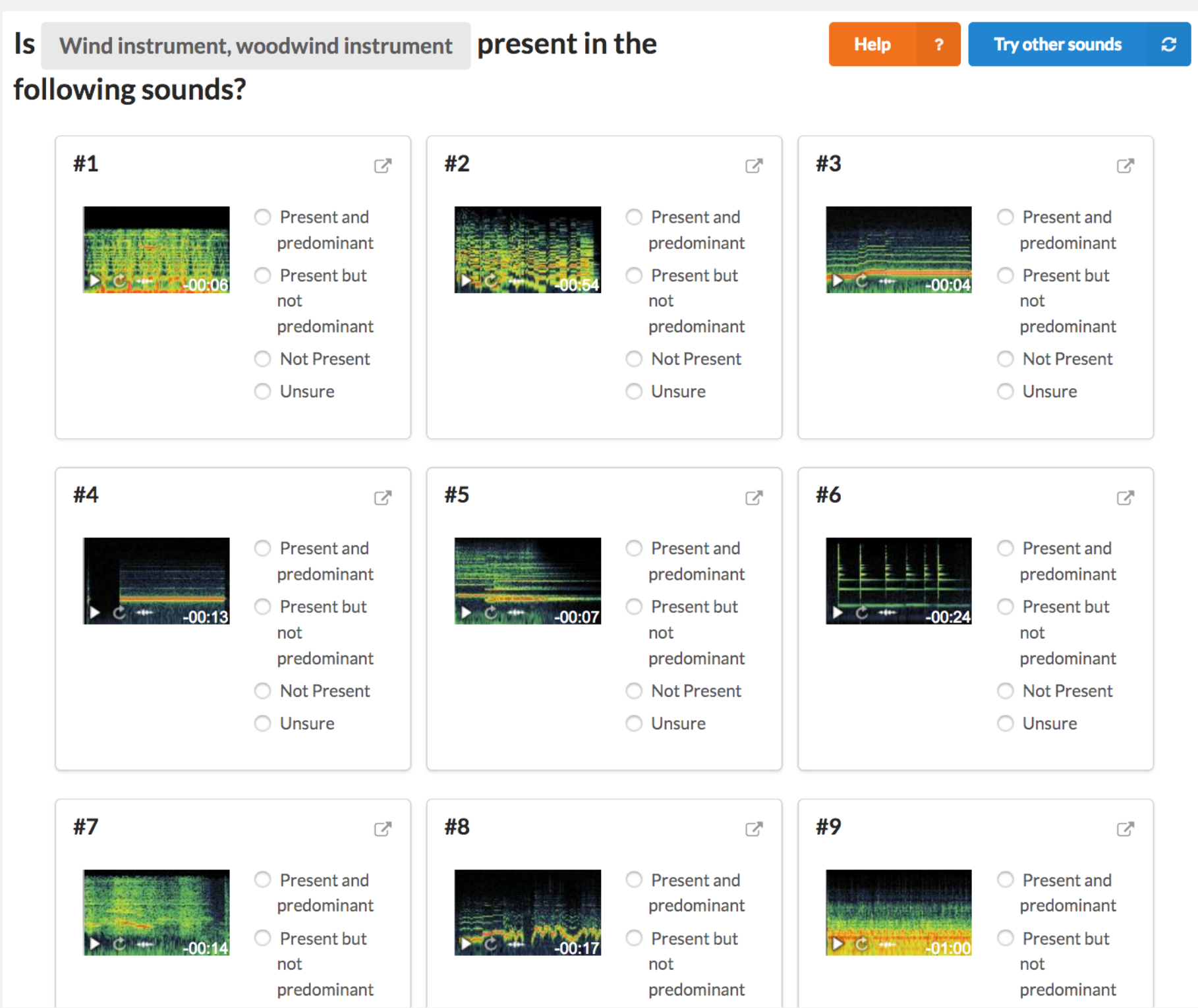
Annotation page for our validation task

Discussion in GitHub repository

Our first dataset: **FSD**

◆ Audio samples from Freesound
◆ Google's AudioSet Ontology [1]
◆ **Data gathering and pre-processing**

◆ General-purpose & large-scale
◆ Audio event recognition

freesound

340k+ sounds → tag matching → 632 classes 300k+ sounds → < 90s → 632 classes 268k sounds → refinement → 398 classes 264k sounds → human validation → **FSD**

AudioSet

feedback → IQA

currently:
23,519 sounds
119h

◆ **Validating Annotations:** presence of sound category
  - Present and Predominant / Present but not Predominant / Not Present / Unsure
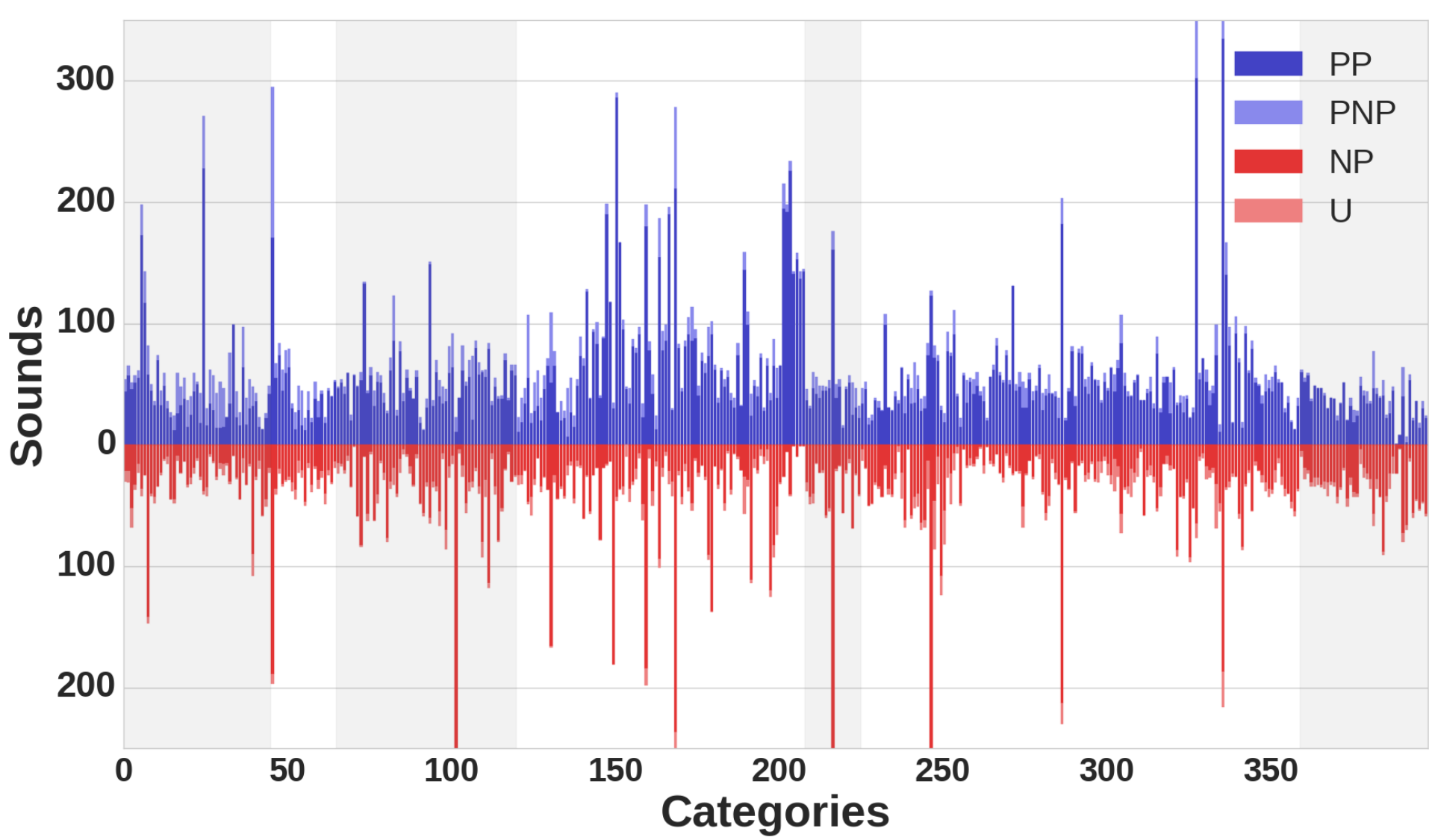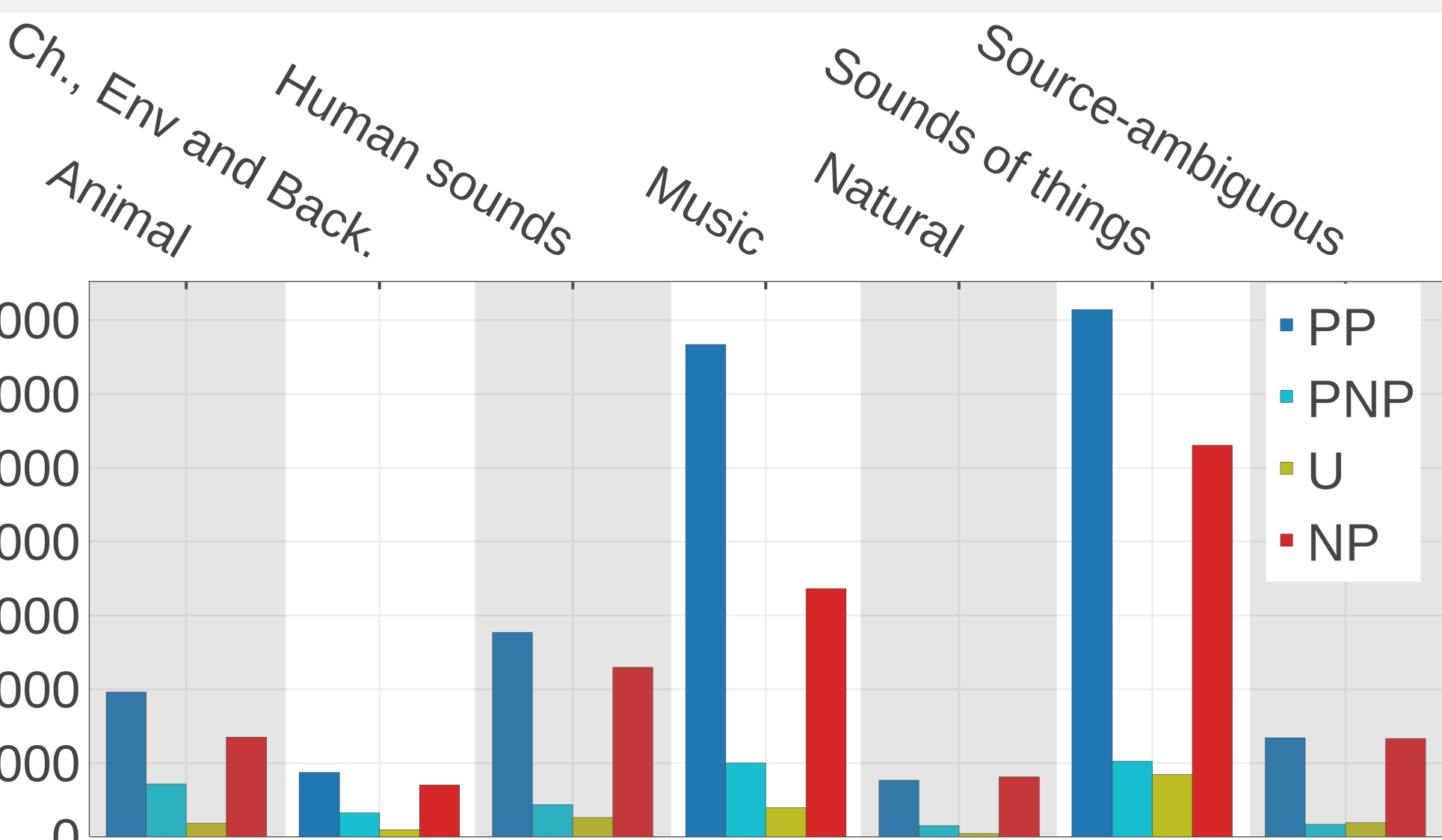
Challenges & future work

◆ **Crowdsource annotations**:
  - quality control
  - annotator agreement
◆ **Mapping** to AudioSet
◆ Towards **complete labeling** of samples
◆ First **FSD release:**
  - data splits & baseline
◆ More **annotation tasks:**
  - timestamps for events
◆ More **functionalities**:
  - report mislabelings & faulty sounds

References & resources

◆ [1] Gemmeke *et al. Audio Set: An ontology and human-labeled dataset for audio events.* ICASSP, 2017
◆ **Platform**: https://datasets.freesound.org/
◆ **GitHub**: https://github.com/MTG/freesound-datasets