LabelBuddy
An Open Source Music and Audio Language Annotation Tagging Tool Using AI Assistance

Ioannis Prokopiou

Orfium | Athens University of Economics and Business

NLP4MusA @ EACL 2026

RABAT, MAROCCO 24-29 March 2026

Funded by the European Union

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

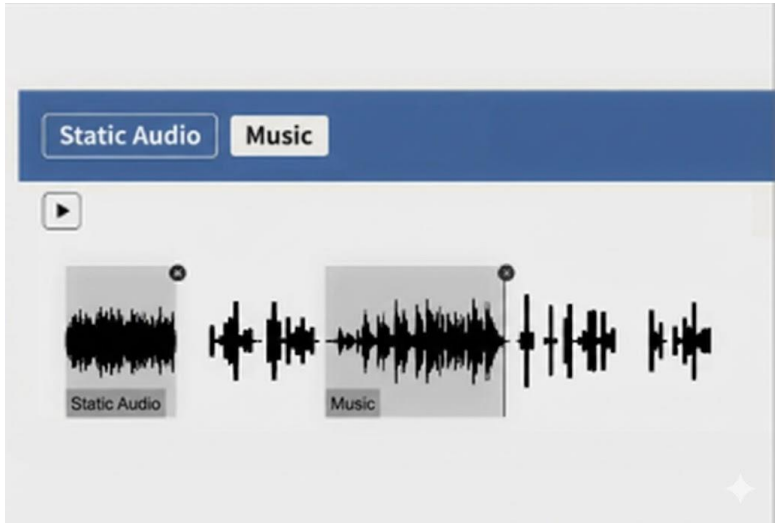
Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

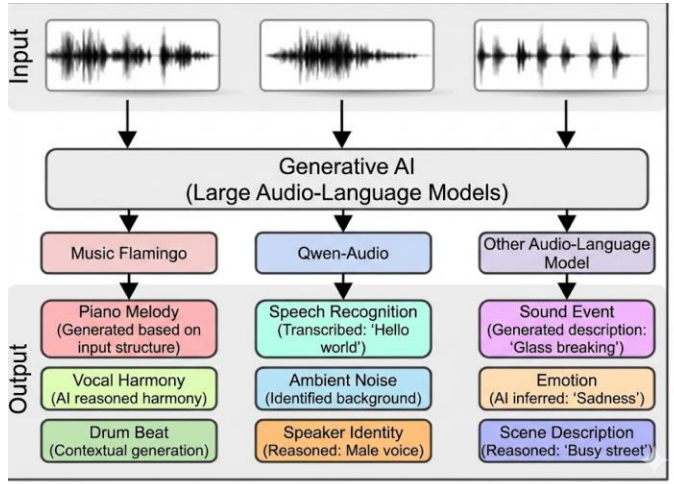
www.aixpert-project.eu

The Shift in MIR & The Problem

Static Tagging



Generative & Reasoning



- **Bottleneck**: Scarcity of open-source infrastructure for subjective nuances.
- Existing tools decouple data curation from subjective evaluation.

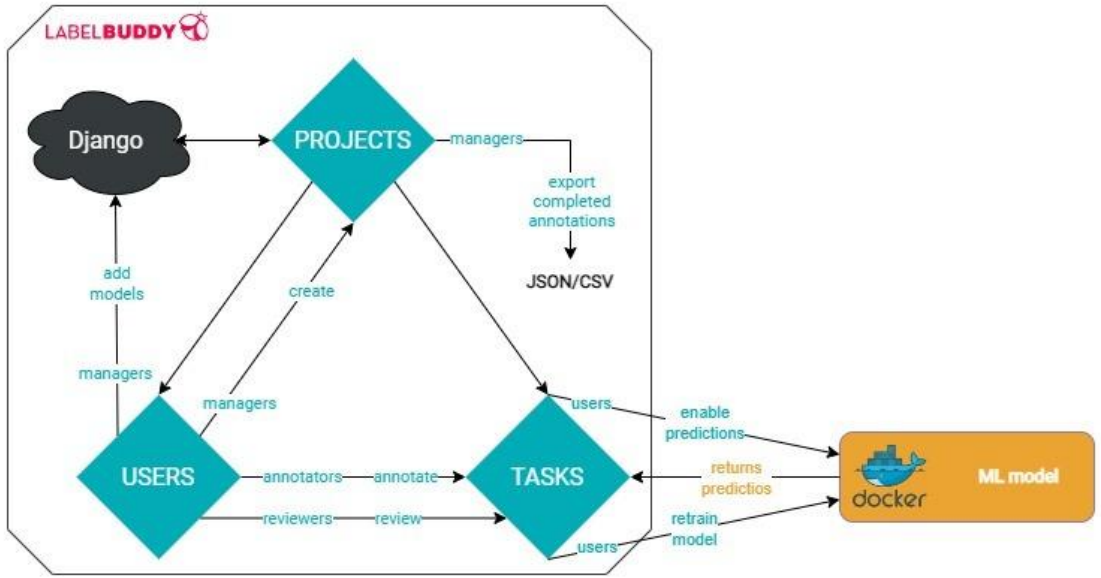
Introducing LabelBuddy

- Three Pillars
  - **Decoupled AI-Assistance**: Containerized inference (Docker).
  - **Collaborative Consensus**: Manager / Annotator / Reviewer roles.
  - **Hybrid Workflow**: Region tagging + Subjective preference.

Table 1: Comparison of LabelBuddy with existing tools.

Tool	Audio Specific	Decoupled AI-Assist	Open Source	Collaboratory Consensus
Audino	✓	-	✓	-
BAT	✓	-	✓	-
Aubio	✓	-	✓	-
Gecko	✓	-	✓	-
Prodigy	-	✓	-	-
Label Studio (CE)	-	✓	✓	-
LabelBuddy	✓	✓	✓	✓

System Architecture



- Django connected to Projects/Users/Tasks, communicating with Docker ML Model
- **Key Tech:** Django (Web Server) + Docker (Inference Engine).
- **Flow:** YAML Config -> Model Container -> JSON Output.

Figure 1: System Architecture Overview: The architecture decouples the Django web server from Docker-ized ML inference.

Workflow & Case Study

- **Example**: Input: Raw Audio + “Instrument Identification” -> AI Output: Piano, Guitar, Drums Segments -> Human Edit: Modification of Regions and verification.
- **Highlights**: AI-generated regions / Editable text / Reviewer Feedback.

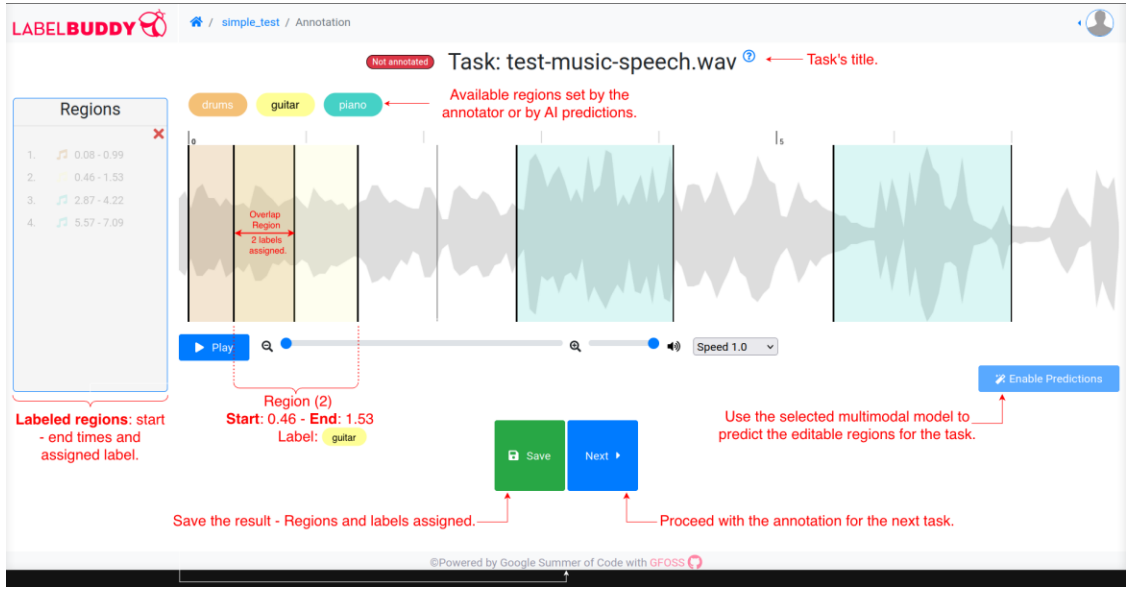


Figure 2: The annotation interface displaying AI-generated predictions as editable waveform regions.

Future Roadmap & Conclusion

- **Future Directions**
  - **Agentic Reasoning:** Conversational assistance (Chain-of-Thought).
  - **RLHF Workbench:** Pairwise Preference (Dataset A vs. Dataset B).
- **Authors**

  - Ioannis Prokopiou (1,2) - gian.prokopiou@aueb.gr
  - Ioannis Sina (3) - sinaioannis@gmail.com
  - Agisilaos Kounelis (3) - agis@ceid.upatras.gr
  - Pantelis Vikatos (2) - pantelis@orfium.com
  - Themos Stafylakis (1,4) - tstafylakis@aueb.gr

GitHub Repo

1 Athens University of Economics and Business
2 Orfium
3 University of Patras
4 Archimedes/Athena R.C.

Our Consortium

ATHENA RESEARCH CENTER

Amsterdam UMC

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

cnrs SCIENCES INFORMATIQUES

Furhat Robotics

INFINITIVITY DESIGN LABS

itml innovation applied

KYKLOS

MARTEL innovate

novelcore RESEARCH & INNOVATION

ORFIUM

PHILIPS

university of groningen

SORBONNE UNIVERSITÉ

UNIVERSITAT DE BARCELONA

VECTOR INSTITUTE | INSTITUT VECTEUR

workable

AIXPERT

Thank you!

aixpert-project.eu

info@aixpert-project.eu

mastodon.social/@AIXPERT_project

/company/aixpert-project/

@AIXPERT_project

Funded by the European Union

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

Funded by the European Union [AIXPERT, 101214389]. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI).