

El Bongosero

A Crowd-sourced Symbolic Dataset of Improvised Hand Percussion Rhythms Paired with Drum Patterns

Nicholas Evans, Behzad Haki, Daniel Gómez-Marín, Sergi Jordà



Universitat
Pompeu Fabra
Barcelona

MTG
Music Technology
Group

1 Introduction

Background

- Traditional drum datasets are typically limited to professional drummers or synthesized data, missing the diversity of general human rhythm perception.
- El Bongosero offers a large-scale, expressive dataset gathered through an interactive installation.

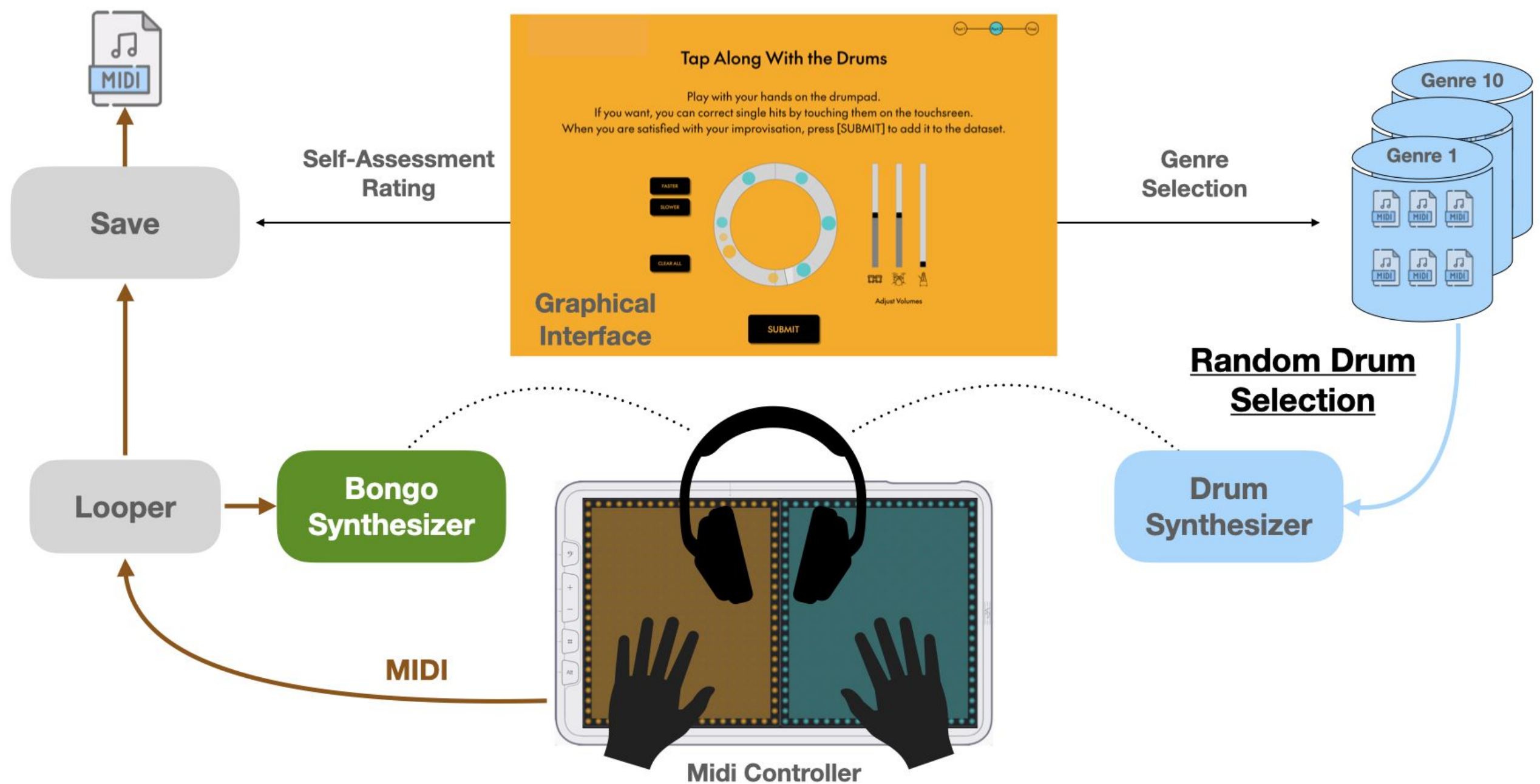
Objective

- To create a dataset that captures improvised drum patterns from a diverse pool of individuals, enriching research on rhythm and music generation.

Dataset	Format			Annotations		
	Audio	Symbolic	Human-Performed	Velocity	Genre	Level of Expertise
ENST	✓	✓	✓			
TMIDT	✓	✓				
ADTOF	✓	✓				
GMD	✓	✓	✓	✓	✓	
MAESTRO	✓	✓	✓	✓		
TapTamDrum	✓	✓	✓	✓		
MAST	✓		✓			
El Bongosero		✓	✓	✓	✓	✓

2 Methodology

- An interactive installation at the Centre de Cultura Contemporània de Barcelona.
- Participants used a MIDI pad to improvise alongside predefined drum patterns.
- Over 3,000 participants contributed, resulting in 6,035 unique sequences.
- Each sequence contains expressive timing, dynamics, and participant musical expertise.
- Participants provided explicit consent for their contributions.
- No personal demographic data was collected, maintaining privacy.



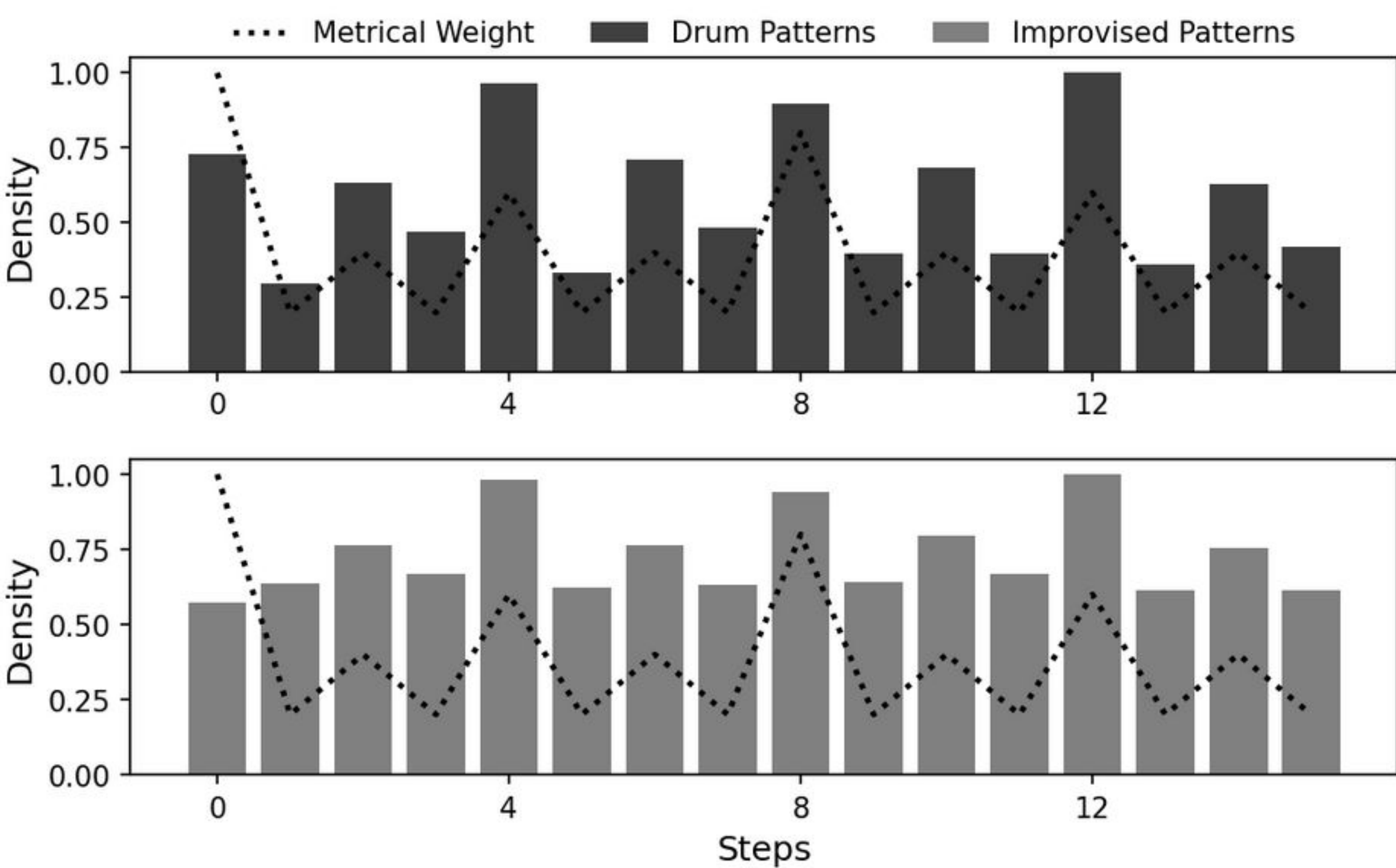
3 Dataset Overview

- 6,035 sequences from 3,184 participants.
- Annotated with timing, dynamics, genre, tempo, expertise level, and enjoyment rating.
- 10 musical genres represented: Afrobeat, Afrocuban, Bossanova, Disco, Electronic, Funk, Hip-hop, Jazz, Rock, Soul.
- Participants ranged from complete novices to experienced musicians.
- Emphasized expressive play rather than strict adherence to patterns

	Level of Musical Expertise				
	1	2	3	4	5
No. participants	392	915	805	594	478
Attempt count	691	1692	1536	1074	1042
Mean no. attempts	1.76	1.85	1.91	1.81	2.18
Unique patterns	433	720	691	562	553

4 Preliminary Analysis

- Evaluated similarity between drum patterns and improvisations, showing a wide range of rhythmic diversity and non-random improvisations.
- Popular genres (Electronic and Rock) had the most improvisations; diversity in expertise levels influenced performance styles.
- Analysis shows participants aligned with metrical expectations, indicating that even novices induce a meter when improvising.

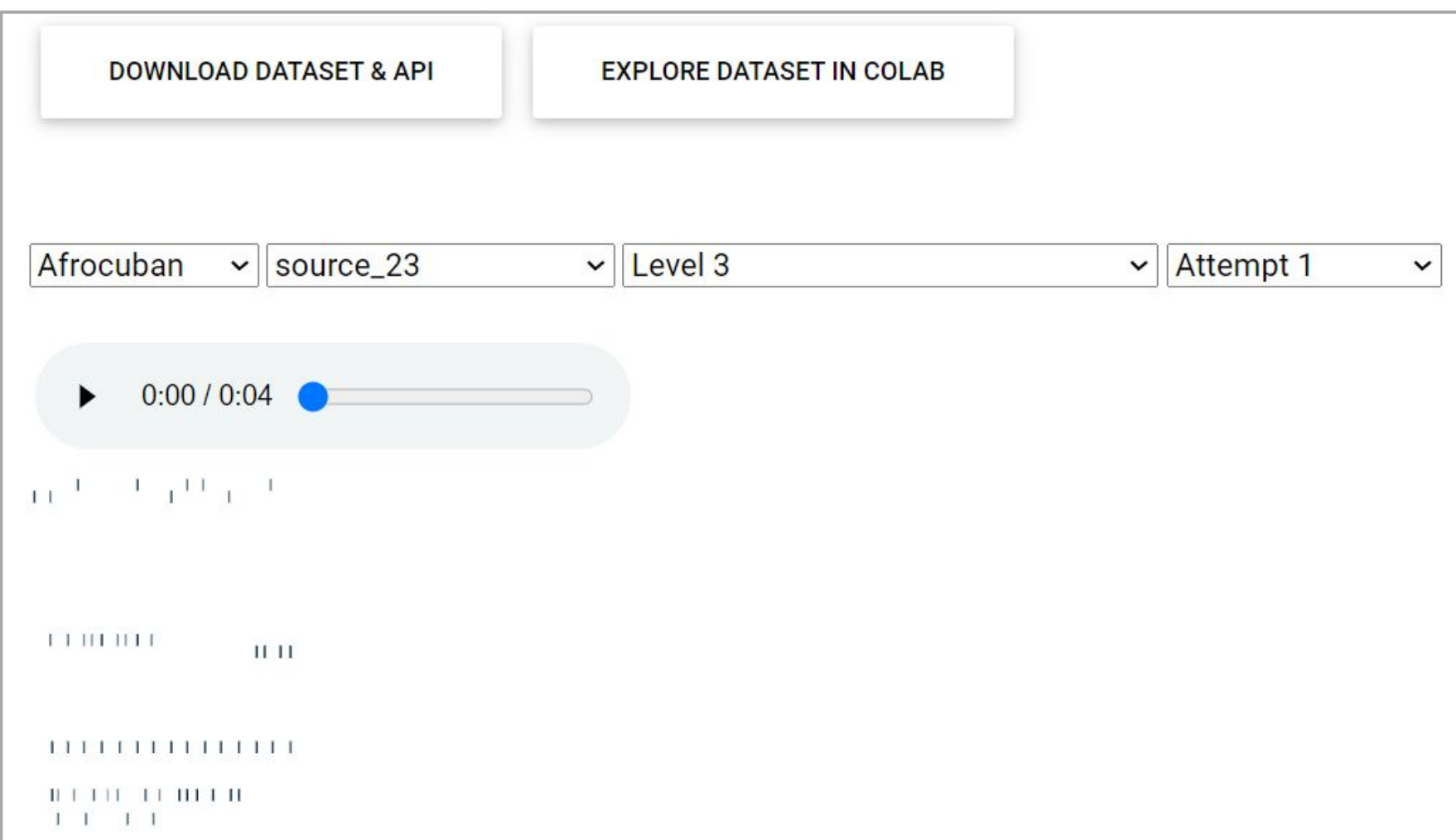


5 Applications

- Improving generative music models to better handle varied skill levels and expressive nuances.
- Research in music education to understand rhythm perception and skill development.

6 Conclusion

- Summary: El Bongosero is a diverse, open-source dataset that captures human improvisational rhythms, filling a gap in current musical datasets.
- Impact: Anticipated to advance music generation technologies and rhythm education studies.
- Todo: Deeper analysis of improvisational behavior across different genres and skill levels.



Dataset and API
<https://elbongosero.github.io>