

# Guitar-Set (Preview): a dataset for guitar transcription and more

centre for digital music



Qingyang Xi\*, Rachel Bittner\*, Johan Pauwels†, Sebastian Ewert†, and Juan P. Bello\*

\*Music and Audio Research Laboratory, New York University

†C4DM, Queen Mary University of London

{tom.xi, rachel.bittner, jpbello}@nyu.edu

{j.pauwels, s.ewert}@qmul.ac.uk

## Summary

- Polyphonic transcription for guitar is hard.
- Labeled data needed for supervised MIR models
- There is no large dataset providing realistic polyphonic guitar recordings and note-level annotations.
- Let's build one!

## Musical Content

30 Structured Excerpts  
~30 sec. each

Rock  
Singer-Songwriter  
Bossa Nova  
Jazz  
Funk

12 Bar Blues  
Autumn Leaves  
Pachelbel Canon

Fast  
Slow

Freestyle  
Playing

## Recording Process



## Challenges

### Note-Level Transcription Accuracy:

- Cross Talk between Strings
- pYin on String Noises

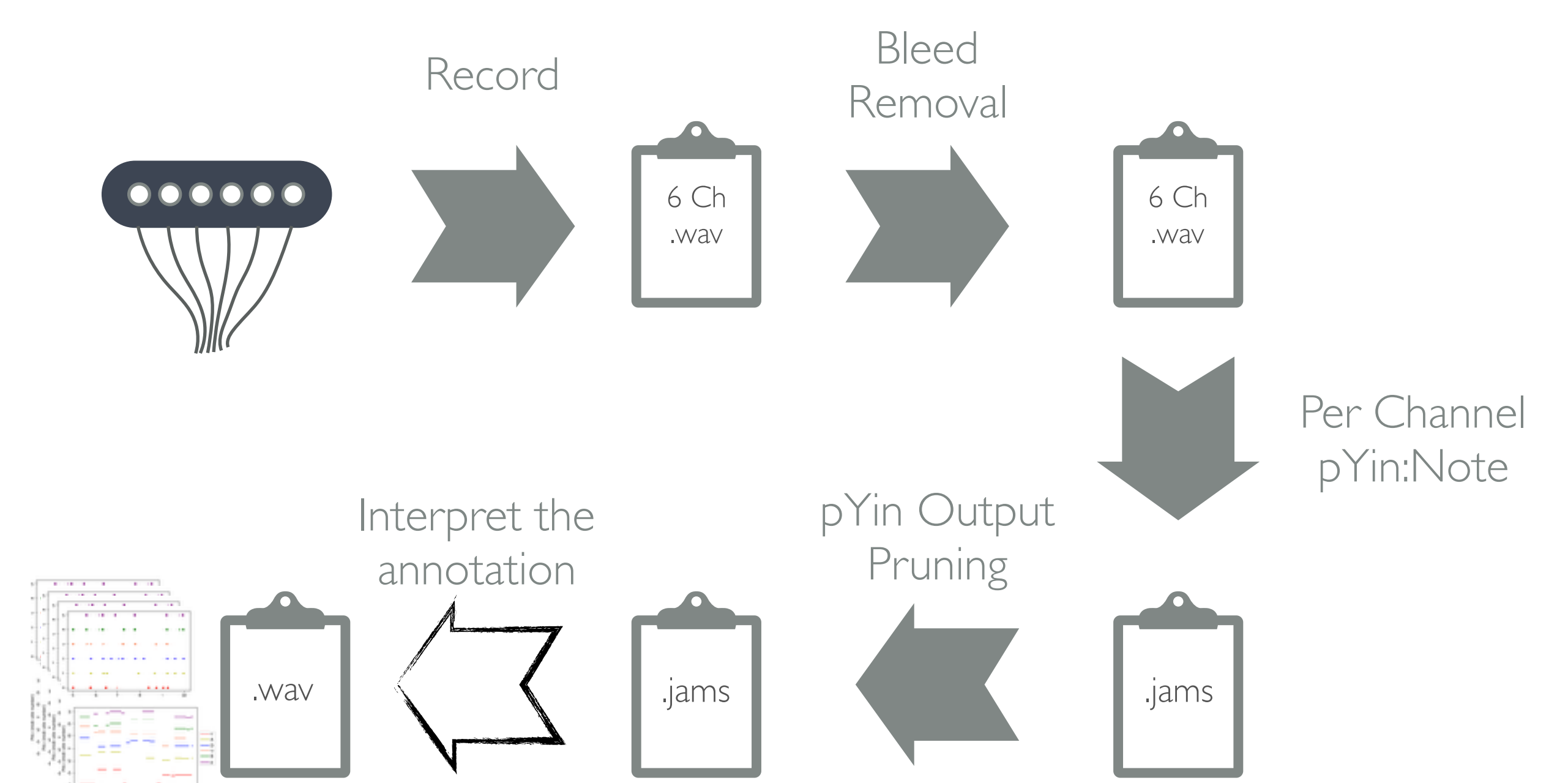
### Intonation:

- Fractional MIDI Note Number

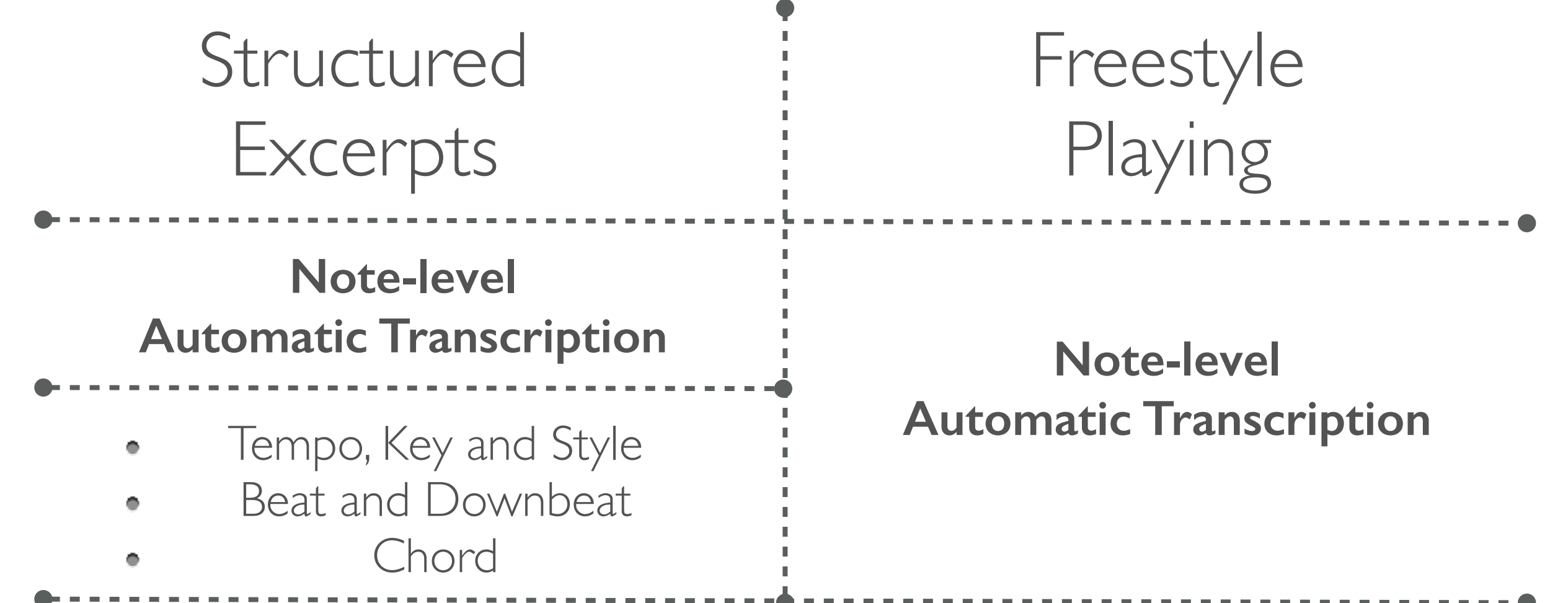
### Playing Techniques:

- Muted or Noise?
- Pull-off Hammer-on
- Harmonics and Sympathetic Resonances

## Annotation Method



## Annotation Types

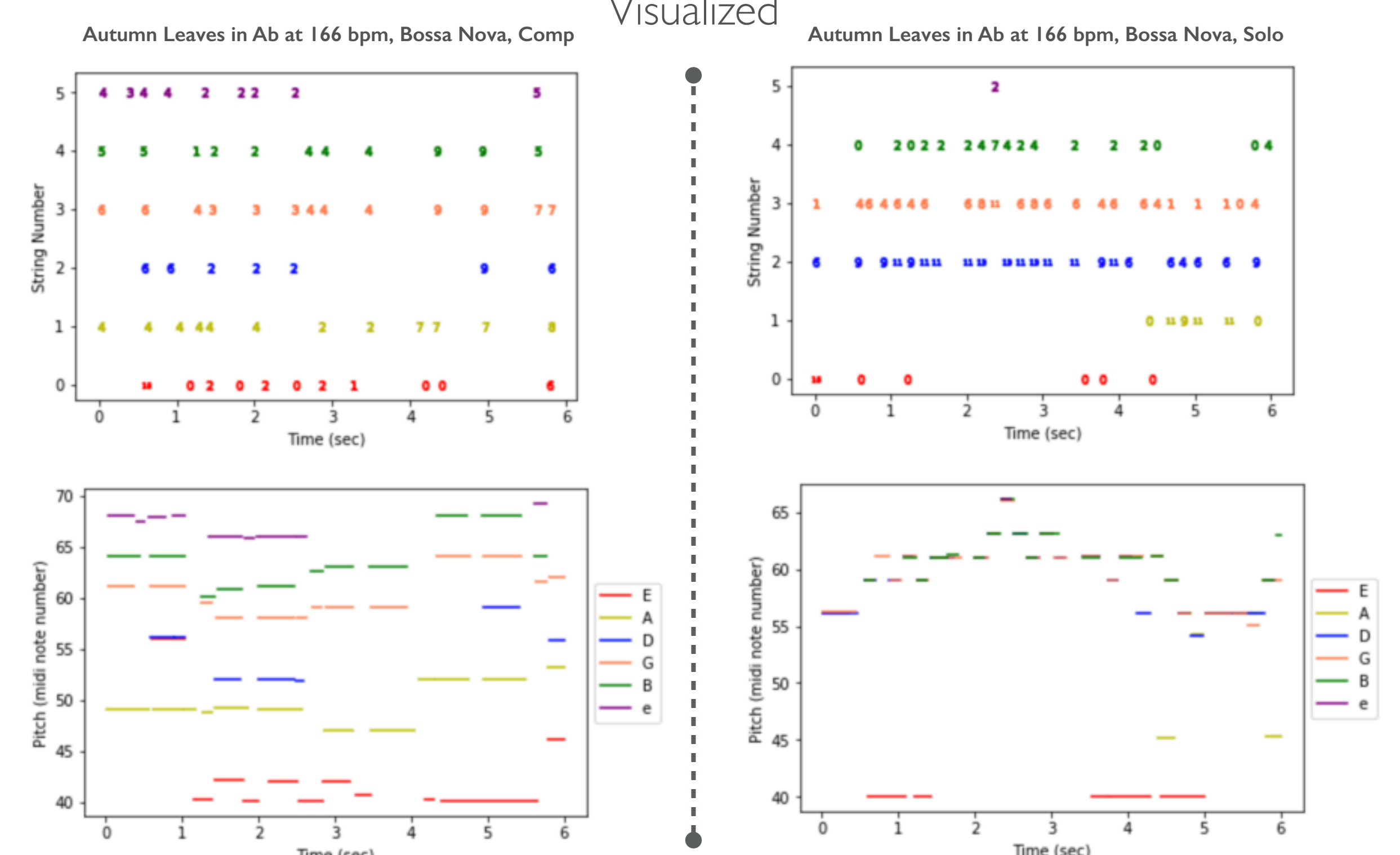


annotate

JAMS file  
JAMS: JSON-based format for storing musical annotations

Interpret

Note-level Automatic Transcription Visualized



WAVE file  
Of the synthesizing of note annotation using midi

WAVE file  
Of the synthesizing of note annotation using midi



NEW YORK UNIVERSITY

