

Preliminary Report on Gathering a Larger Annotated Dataset for Pattern Discovery Tasks

Klara Žnideršič, Iris Y. Ren, Anja Volk, Matevž Pesek

Pattern discovery

- Since ~2013 @ MIREX
- Johannes Kepler University Patterns Development/Test Databases
 - 5 songs in each dataset
- Four subtasks
 - symPoly, symMono (converted)
 - audPoly, audMono
- Small group of participants

A *Tempo di Menuetto* [$\text{♩} = 120$]

0 (36) 3 (39) 6 (42) 9 (45) 12 (48) 15 (51)

7 18 (54) 21 (57) 24 (60) 27 (63) 30 (66) 33 (69)

B

P_1	P_2	R_1	R_2		P_3	P_4	R_3	R_4	
-------	-------	-------	-------	--	-------	-------	-------	-------	--

C

Q_1		Q_2	
-------	--	-------	--

D

S_1		S_2
-------	--	-------

0 6 12 18 24 30 36 42 48 54 60 66

Following years

- Still a small group of participants
- Limitations of the JKU-PDD (5 songs)
- Interest remains, but the task is somehow limited
- spin-off task -> Patterns for prediction
- Two subtasks
 - (explicit) - matching algo's prediction with original
 - (implicit) – providing probability of the given continuation being genuine

Sub code	Submission name	Abstract	Contributors
Task Version	symPoly		
DM1	SIATECCompress-TLF1	PDF D	David Meredith E
DM2	SIATECCompress-TLP	PDF D	David Meredith E
DM3	SIATECCompress-TLR	PDF D	David Meredith E
Task Version	symMono		
DM1	SIATECCompress-TLF1	PDF D	David Meredith E
DM2	SIATECCompress-TLP	PDF D	David Meredith E
DM3	SIATECCompress-TLR	PDF D	David Meredith E
IR1	mypattern	PDF D	Iris YuPing Ren E
PLM1	SYMCHM	PDF D	Matevž Pesek E , Urša Medvešek, Aleš Leonardis E , Matija Marolt E
VM1'14	VM1	PDF D	Gissel Velarde E , David Meredith E
VM2'14	VM2	PDF D	Gissel Velarde E , David Meredith E

Sub code	Submission name	Abstract	Contributors
Task Version	symMono		
NF1	MotivesExtractor	PDF D	Oriol Nieto E , Morwaread Farbood E
OL1	PatMinr	PDF D	Olivier Lartillot E
VM1	VM1	PDF D	Gissel Velarde E , David Meredith E
VM2	VM2	PDF D	Gissel Velarde E , David Meredith E
NF1'13	motives_mono	PDF D	Oriol Nieto E , Morwaread Farbood E
DM10'13	SIATECCompressSegment	PDF D	David Meredith E
Task Version	symPoly		
NF1	MotivesExtractor	PDF D	Oriol Nieto E , Morwaread Farbood E
NF2'13	motives_poly	PDF D	Oriol Nieto E , Morwaread Farbood E
DM10'13	SIATECCompressSegment	PDF D	David Meredith E
Task Version	audMono		
NF1	MotivesExtractor	PDF D	Oriol Nieto E , Morwaread Farbood E
NF3'13	motives_audio_mono	PDF D	Oriol Nieto E , Morwaread Farbood E
Task Version	audPoly		
NF1	MotivesExtractor	PDF D	Oriol Nieto E , Morwaread Farbood E
NF4'13	motives_audio_poly	PDF D	Oriol Nieto E , Morwaread Farbood E

Sub code	Submission name	Abstract	Contributors
Task Version	symMono		
PLM1	SYMCHM	PDF D	Matevž Pesek E , Urša Medvešek, Aleš Leonardis E , Matija Marolt E
OL1'14	PatMinr	PDF D	Olivier Lartillot E
VM2'14	VM2	PDF D	Gissel Velarde E , David Meredith E
Task Version	audMono		
WHD1	VMO Motif Discovery	PDF D	Cheng-i Wang E , Jennifer Hsu E , Shlomo Dubnov E
WDH1	VMO Motif Discovery FML	PDF D	Cheng-i Wang E , Jennifer Hsu E , Shlomo Dubnov E
NF1'14	MotivesExtractor	PDF D	Oriol Nieto E , Morwaread Farbood E
Task Version	audPoly		
WHD1	VMO Motif Discovery	PDF D	Cheng-i Wang E , Jennifer Hsu E , Shlomo Dubnov E
WDH1	VMO Motif Discovery FML	PDF D	Cheng-i Wang E , Jennifer Hsu E , Shlomo Dubnov E
NF1'14	MotivesExtractor	PDF D	Oriol Nieto E , Morwaread Farbood E

Rebooting the pattern discovery

- Meetup at ISMIR 2019
- Mostly researchers who submitted to MIREX tasks
- High interest remained
- *Patterns for prediction* task



What happened?

- Covid-19
 - Less tasks in 2020
 - No Patterns for prediction or Pattern discovery tasks in 2021
 - No MIREX tasks in 2022
 - Stephen Downie – interest, finances, infrastructure issues

2021:MIREX2020 Results

Results by Task (More results are coming)

- 2021: Automatic Lyrics Transcription Results
- 2021: Set List Identification

2020:MIREX2020 Results

Overall Results Poster

Coming soon

Results by Task (More results are coming)

- [Audio Fingerprinting Results](#)
- [Audio Melody Extraction Results](#)
 - [ADC04 Dataset](#)
 - [MIREX05 Dataset](#)
 - [INDIAN08 Dataset](#)
 - [MIREX09 0dB Dataset](#)
 - [MIREX09 -5dB Dataset](#)
 - [MIREX09 +5dB Dataset](#)
 - [ORCHSET15 Dataset](#)
- [Patterns for Prediction Results](#)
- [Audio Cover Song Identification Results](#)
- [Train-Test Task Set](#)
 - [Audio Music Mood Classification Results](#)
 - [Audio KPOP Mood \(Annotated by Korean Annotators\) Classification Results](#)
 - [Audio KPOP Mood \(Annotated by American Annotators\) Classification Results](#)
 - [Audio Mixed Popular Genre Classification Results](#)
 - [Audio KPOP Genre \(Annotated by Korean Annotators\) Classification Results](#)
 - [Audio KPOP Genre \(Annotated by American Annotators\) Classification Results](#)
- [Automatic Lyrics-to-Audio Alignment Results](#)
- [Lyrics Transcription Results](#)
- [Audio Key Detection](#)
- [Audio Chord Estimation](#)
 - [Isophonics2009 Dataset](#)
 - [Billboard2012 Dataset](#)
 - [Billboard2013 Dataset](#)
 - [JayChou29 Dataset](#)
 - [RobbieWilliams Dataset](#)
 - [RWC-Popular Dataset](#)
 - [USPOP2002Chords Dataset](#)
 - [CASD Dataset](#)
- [Singing Transcription from Polyphonic Music Results](#)

results by year

- [MIREX 2021 Results](#)
- [MIREX 2020 Results](#)
- [MIREX 2019 Results](#)
- [MIREX 2018 Results](#)
- [MIREX 2017 Results](#)
- [MIREX 2016 Results](#)
- [MIREX 2015 Results](#)
- [MIREX 2014 Results](#)
- [MIREX 2013 Results](#)
- [MIREX 2012 Results](#)
- [MIREX 2011 Results](#)
- [MIREX 2010 Results](#)
- [MIREX 2009 Results](#)
- [MIREX 2008 Results](#)
- [MIREX 2007 Results](#)
- [MIREX 2006 Results](#)
- [MIREX 2005 Results](#)

However ...

- Activity (and hope) remains!
 - Submissions to ISMIR 2022/23
 - SIATEC-C: Computationally efficient repeated pattern discovery in polyphonic music (Björklund)
 - Active publications by “members” of the former Pattern discovery task
 - Understanding and Compressing Music with Maximal Transformable Patterns (Meredith) - 2021
 - **Exploring annotations for musical pattern discovery gathered with digital annotation tools** (Tomašević et al.) – 2021
 - A Computational Evaluation of Musical Pattern Discovery Algorithms (Ren et al.) – 2020
 - **A Comparison of Human and Computational Melody Prediction Through Familiarity and Expertise** (Pesek et al.) - 2020

Similar topics - prediction

- Predicting melody continuation
 - European and Chinese folk song datasets
 - Complete and incomplete sequences
 - Predicting the next note in a melodic sequence (57 participants)
 - Using of SymCHM model for prediction
 - Comparing to human participants (musicians/non-musicians)

- Residual results
 - Slovenian translation and validation of MSI instrument (231 participants)



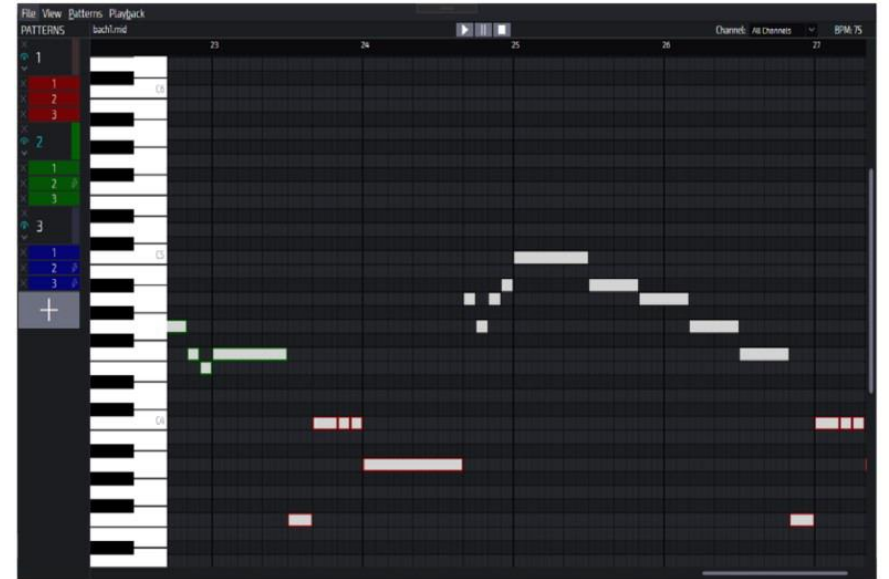
Model	Dataset	
	European	Chinese
SymCHM-eu	0.720	0.491
SymCHM-cn	0.635	0.520
Adjusted I-R model	0.648	0.490

	European excerpts			Chinese excerpts		
	All (%)	Complete (%)	Incomplete (%)	All (%)	Complete (%)	Incomplete (%)
All participants	58	63	42	34	34	35
Musicians	68	74	47	39	37	44
Non-musicians	41	43	33	25	27	21
baseline (all notes)	4	4	4	4	4	4
baseline (scale notes)	6.7	6.7	6.7	9.1	9.1	9.1
SymCHM-eu	60	73	20	30	33	20
SymCHM-cn	45	53	20	30	40	0
Adjusted I-R model	50	60	20	35	25	40

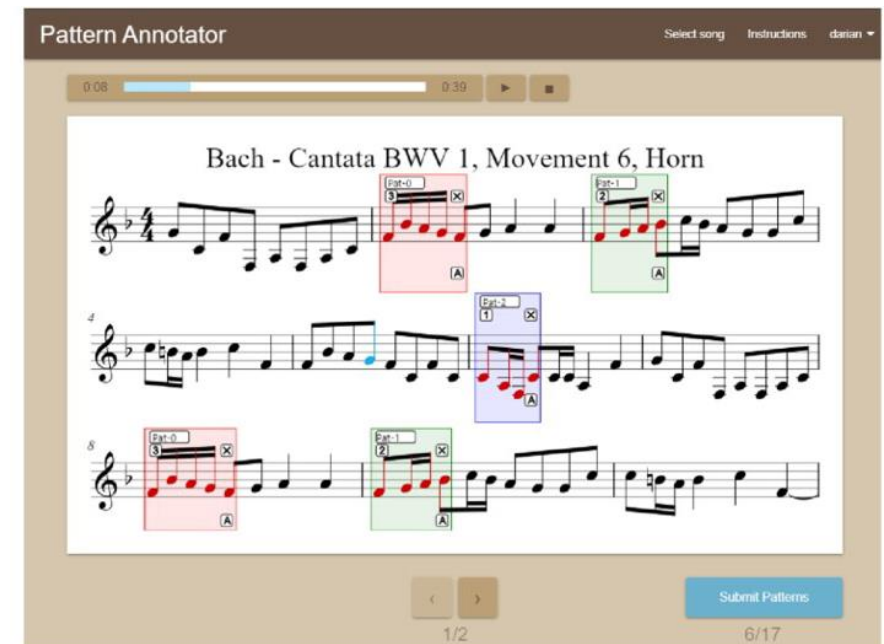
The table also includes a baseline, taking into account all possible 25 semitones, which could be picked from the interface, and a scale notes baseline, which includes 15 semitones for major/minor scales (European excerpts), and 11 semitones for pentatonic scale (Chinese excerpts), both within \pm one octave.

Similar topics - annotations

- Brainstorming since 2019
 - Collecting multiple annotations -> inter-annotator agreement
 - Two tools (ANOMIC and PAF)
 - Nieto-Farbood (ANOMIC) Dataset
 - Bach, Beethoven, Haydn, Mozart (6 pcs)
 - 13 (PAF) + 26 (Anomic) students
 - 4 Musicology (MU), 3 Theory and composition (TC), 6 Pedagogy (PE)

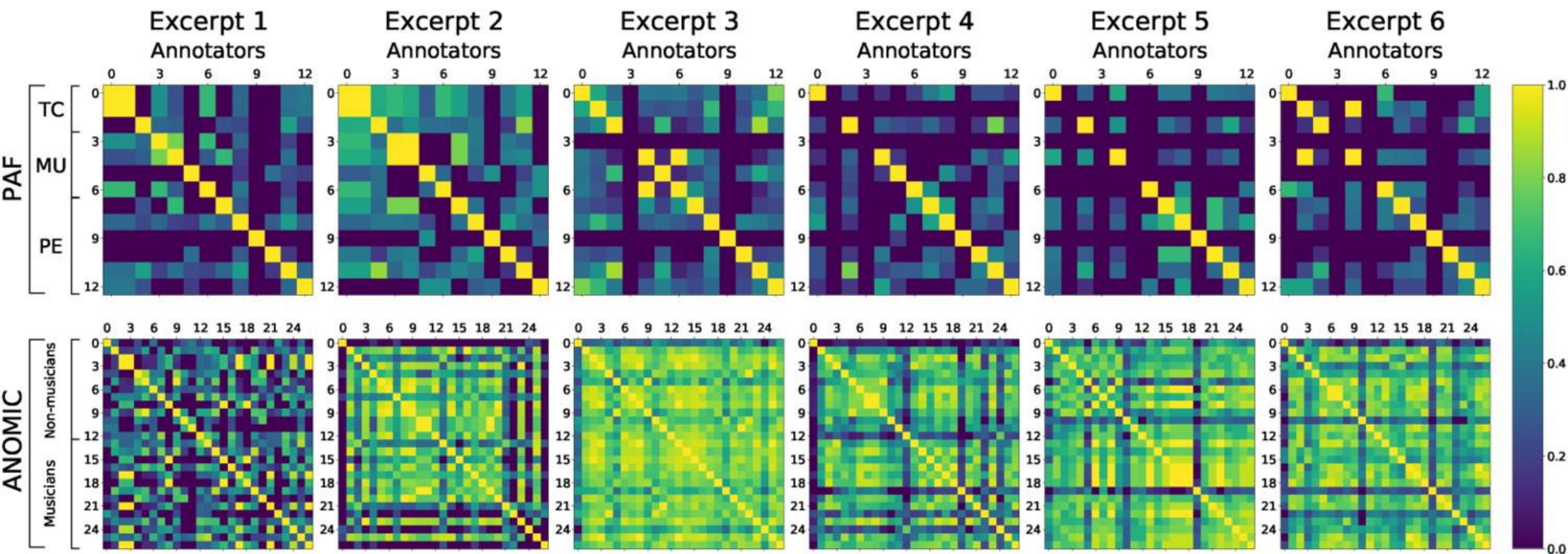


(a) ANOMIC interface



(b) PAF interface

Results

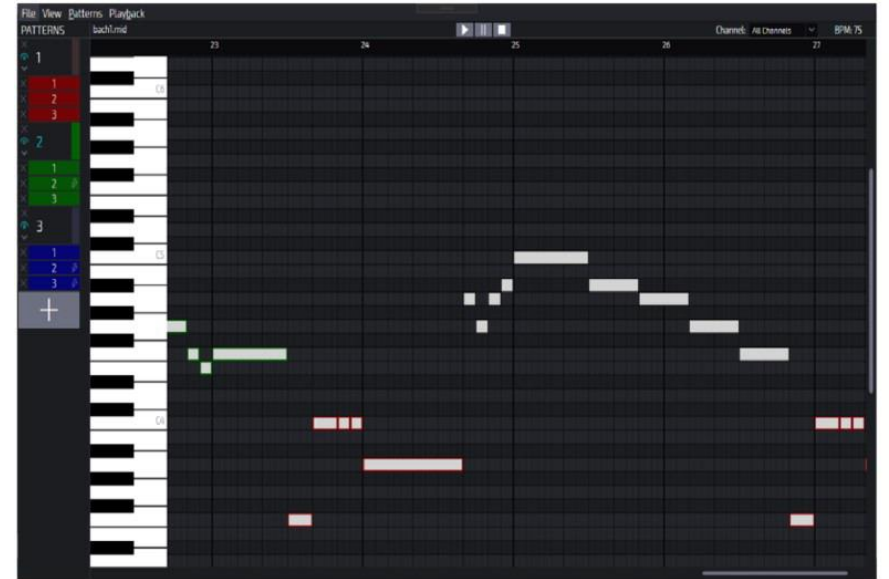


What the #3&% is a pattern?

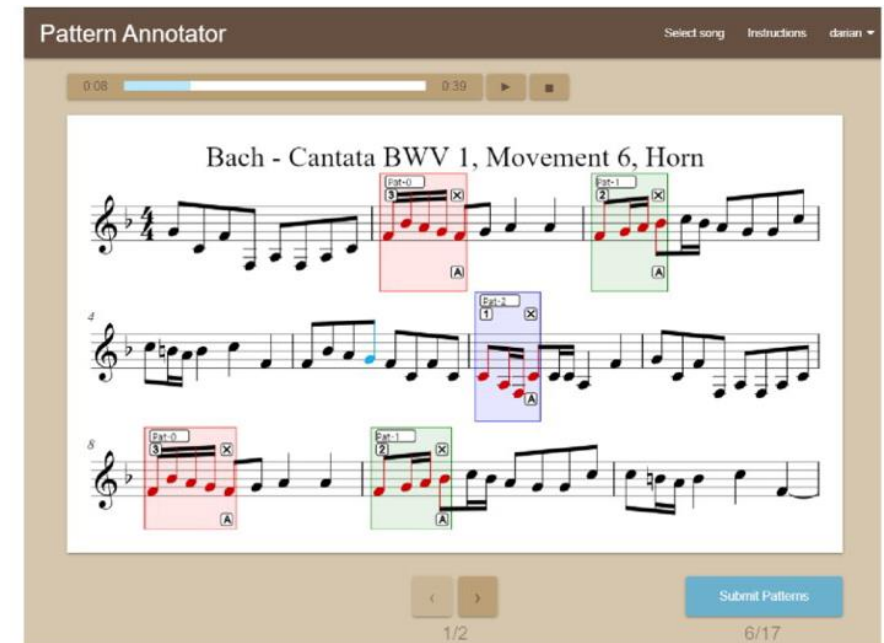
- We all know there are patterns (or do we?)
 - What defines them? Why do they differ between different experts?
- How is the expert's background influencing their perception of a pattern?
 - Expertise, familiarity, instrument, other social aspects
- Should we consider all patterns (simultaneously)?
- What goal(s) should the algorithms pursue?

Comparing differences

- Different interfaces
 - Music notations vs. piano roll
 - Standalone vs. web
- Different student groups
 - Different background, similar age (i.e. experience)
- Different annotations
 - Pattern importance, naming



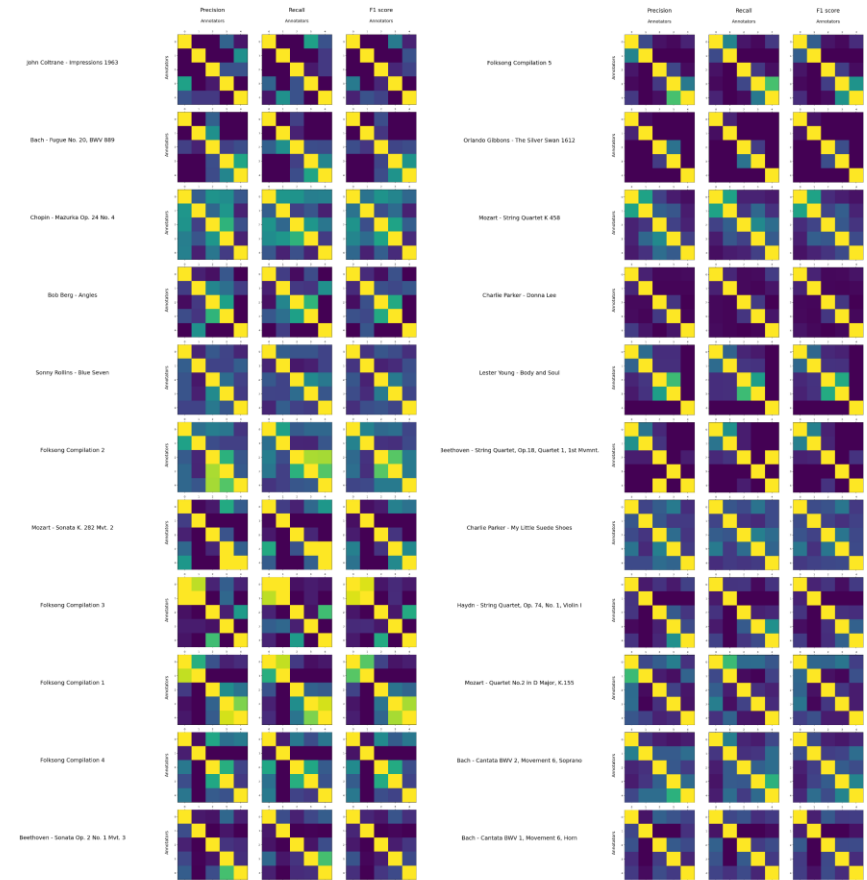
(a) ANOMIC interface



(b) PAF interface

What's with the inter-annotator agreement?

- Should we consider all results as reference annotations concurrently?
- Should we compare results to annotator groups (education, instrument, years of experience/profession)?
- How do we take these findings into consideration?



The idea

1. We need a bigger dataset
 - a. More songs
 - b. More diverse
2. We need more annotators
 1. Big (enough) group/s
 2. Willpower (or financial motivation) to annotate that bigger dataset

The dataset

- Retaining the comparability with the previous task
 - JKU-PDD - 5 songs
 - Excerpts (used for ANOMIC/PAF comparison) – 6 songs
- Novelty
 - Jazz excerpts – 6 songs
 - Folk songs – 5 songs (each includes 6 variants within tune family, concatenated)
- *Potentially missing*
 - *Modern popular genres*
 - *(Maybe) more constrained pattern definition*

The annotators

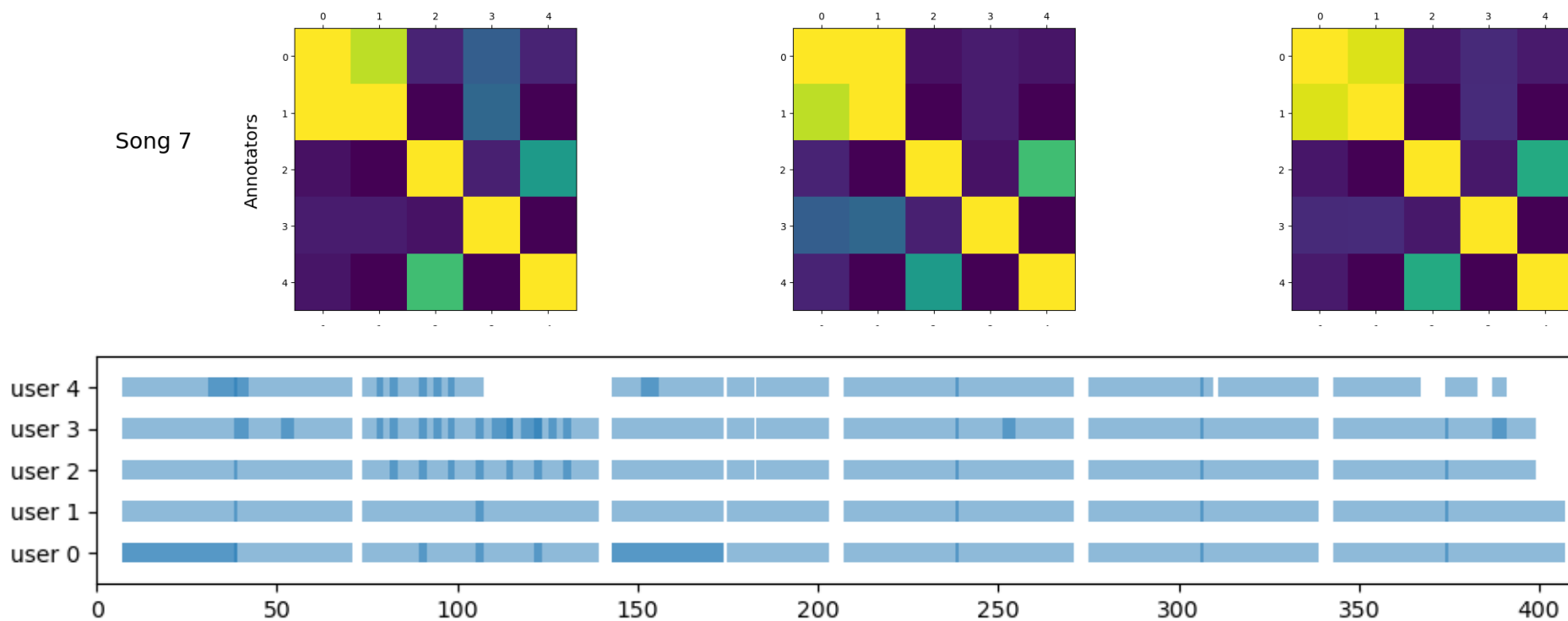
- Five annotators (2F, 3M; 20-25 years old)
- Musicology masters programme
- 14 - 30 days for annotations per annotator (may – dec 2022)
- Financial motivation (student work)

Preliminary results

- 4026 patterns/occurrences in 22 songs
- 2.1 pattern occurrences per song
 - Max per annotator: 36, 19, 19, 6, 16
- Average pattern duration: 8.18 notes

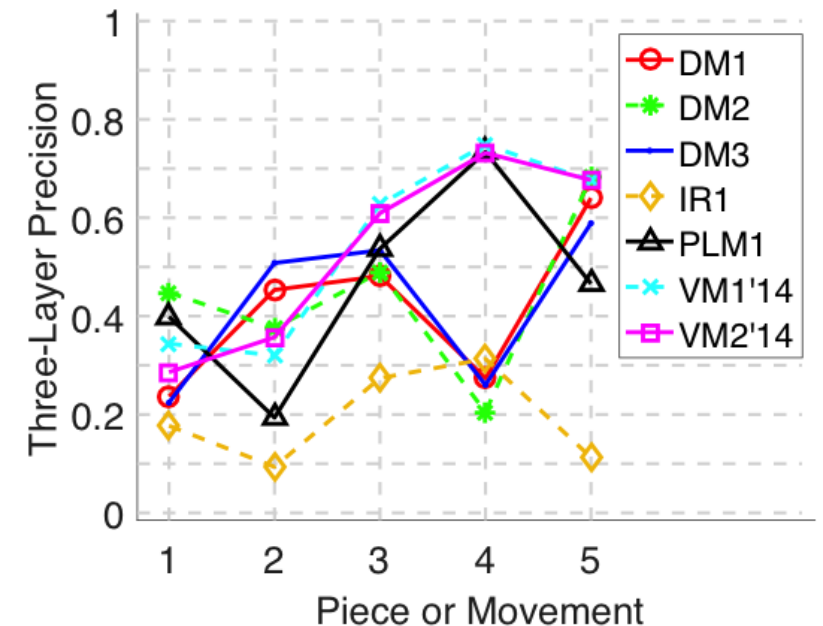
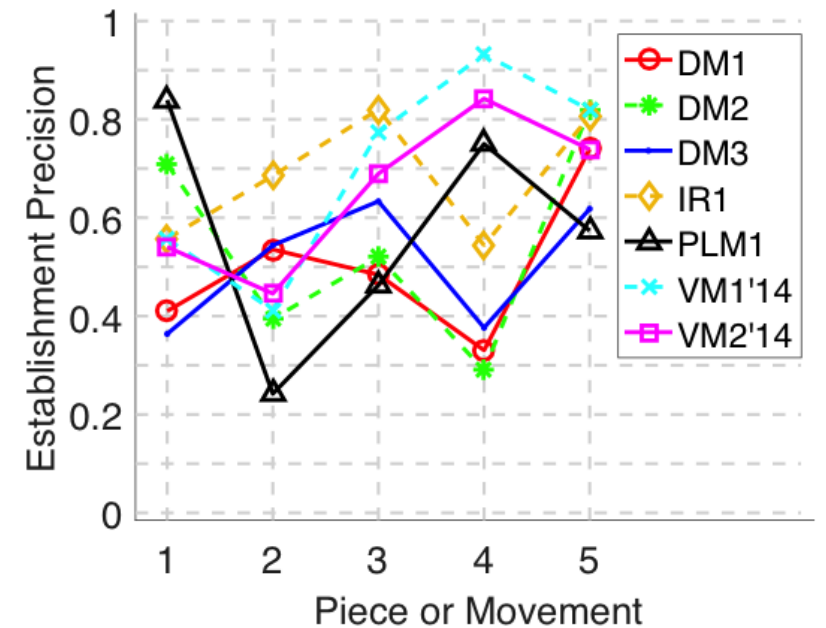
- Pattern types:
 - Transformations - prime forms, inversions, retrograde, retrograde inversions (annotated explicitly)
 - Sub-patterns (explicitly marked)

Inter-annotator agreement



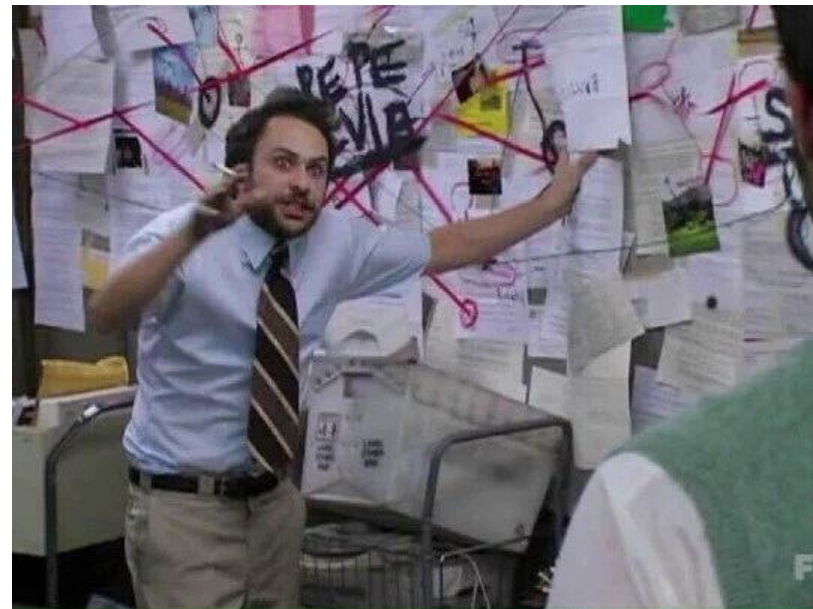
Immediate steps

- Cleaning up the data
 - Connect differently-named patterns as single entities (complex issue, i.e. subpatterns, types etc)
- Inter-annotator agreement analysis
 - Outlier identification, comparison of previous pattern annotations of included sub-sets
 - Identifying the “common ground” within a (relatively) non-diverse group
- Using the data to evaluate (existing) approaches to pattern discovery
 - Contact MIREX task authors and evaluate their approach using the *standard* metrics



Further steps

- Figuring out the adjusted metrics - how to include dis/agreement?
 - Fitting the algorithm's output to one annotator group (experience, background)
- Getting another group of annotators
 - Financial limitations and time consumption
- Further enlarging the dataset
 - Time/finance, access to groups ...



Looking forward to share the dataset and (re)evaluate algorithms!

Klara Žnideršič, Iris Y. Ren, Anja Volk, Matevž Pesek
matevz.pesek@fri.uni-lj.si / matevz.pesek@gmail.com