

Integration of artificial intelligence technologies in the digitization of cultural heritage

Student: Linda Luīze Barbare

PhD supervisor: Dr. sc. comp. Kārlis Freivalds

Working under the project: Spoken Songs: The Algorithms of Composition and Improvisation No. Izp-2025/1-0252



Analysis of Latvian spoken songs using artificial intelligence

- Date back to pre-18th century
- Characterized by narrow melodic ranges (3–5 pitches)
- The research novelty is in the application of optical music recognition (OMR), optical character recognition (OCR), and artificial intelligence (AI) methods in the digitization and analysis of spoken songs.
- In the project an interdisciplinary team of researchers will:
 - encode melodies and texts
 - conduct algorithmic analysis
 - conduct qualitative case studies
 - identify and map compositional and improvisational skills and techniques, focusing on the interaction between text and melody and on local specificities.



Where are we at right now?

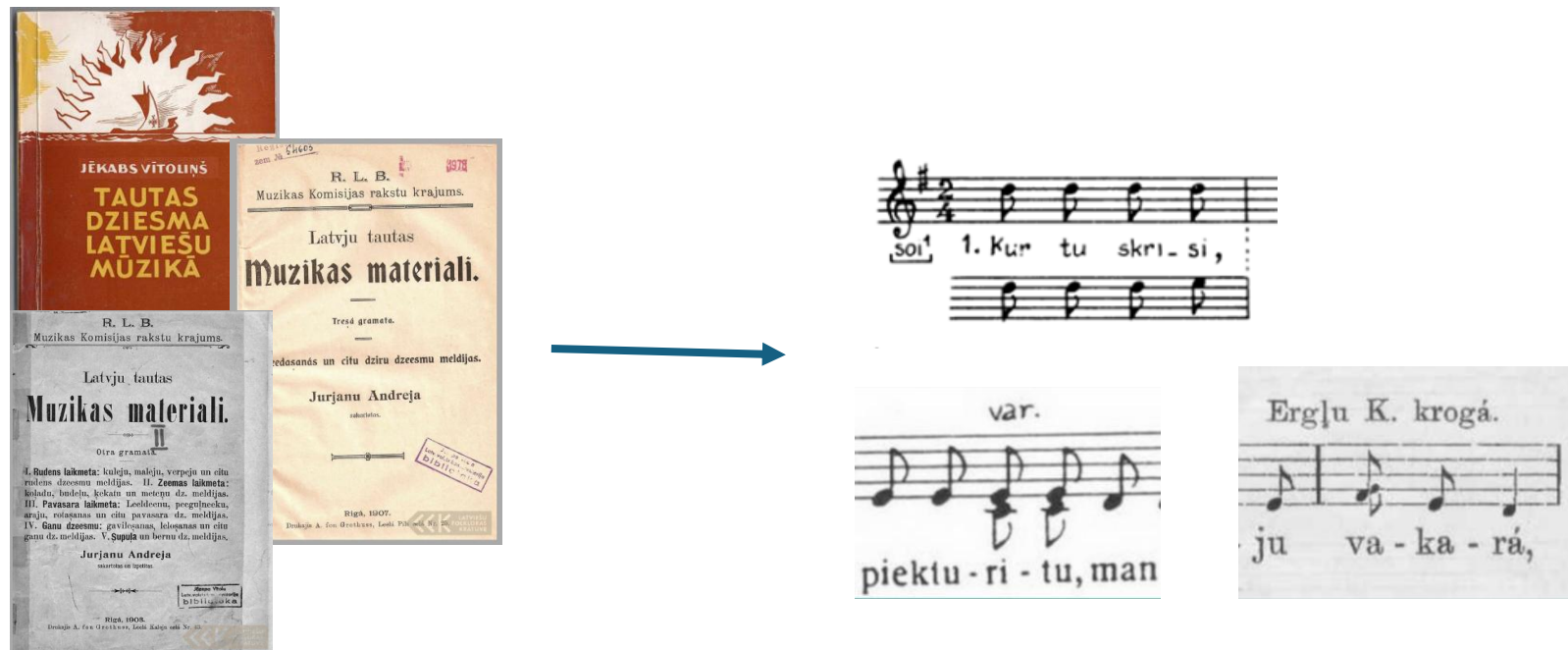
In order to analyse spoken song melodies and texts together the first step is to digitalise everything which includes:

- Melodic digitization (using OMR and OCR)
- Textual variation recovery (using LLM)

Note: dainuskapis.lv has only 222 079 out of 3 000 000 dainas digitalised.

Melodic digitization (using OMR and OCR)

- The dataset contains ~1000 samples of sheetmusic.
- There are 3 sources which means 3 different ways the composers have notated a melodic variant.
- The challenge: automate scanned sheetmusic transformation into musicXML format.



Testing different OMR methods



Variant a of the song «Kur tu skrīsi sila pele» from the Jēkabs Vītoliņš book «Latviešu tautas mūzika. "Gadskārtu ieražu dziesmas" (1973)»

15 Bronislava Justa Rēzeknes Varakļānos, dzim. 1906; pier. E. Melngailis 1941. FS 1045, 9398.

1. Kur tu skrī-si, sy-la pē-lē? Rū-to, rū-tō!

1. Uz zēr-de-ņu zēr-ņu zog-tu, Rū-tō, rū-tō!

2. Kai tur teva nanūsissi?
— As ībēgšu aleņāi.

3. Kai tu tur i nasasalsi?
— As sakuršu kaulu guni.

4. Kai tu tur i nasadegsi?
— Sasaspraužšu odotōsi.

Variant b of the song «Kur tu skrīsi sila pele» from the Emīls Melngailis book «Latviešu mūzikas folkloras materiāli, 2. sējums "Maliēna" (1952)»

	OMR		OCR	
	a	b	a	b
SoundSlice	100%	99%	64%	64-48%
Gemini	57%	80%	100%	80-92%
Audiverius	87%	24%	0%	0%

Testing different OMR methods: SoundSlice



Variant a of the song «Kur tu skrīsi sila pele» from the Jēkabs Vītoliņš book «Latviešu tautas mūzika. "Gadskārtu ieražu dziesmas" (1973)»



Variant a processed through SoundSlice

OMR								
Category	Total Items in the sample	True positives	False negatives (missing items)	False positives (extra items)	F1	Weights	Weighted accuracy	Percent
Pitch	24	24	0	0	1.00	0.46	0.46	46%
Rhythm	24	24	0	0	1.00	0.46	0.46	46%
Measures	2	2	0	0	1.00	0.04	0.04	4%
Key Signature	1	1	0	0	1.00	0.02	0.02	2%
Time Signature	1	1	0	0	1.00	0.02	0.02	2%
TOTAL	52	52	0	0	5.00	1.00	1.00	100%
OCR								
Category	Total Items in the sample	True positives	False negatives (missing items)	False positives (extra items)	F1	Weights	Weighted accuracy	Percent
Metadata	1	0	1	0	0.00	0.04	0	0.00%
Syllables	24	16	8	8	0.67	0.96	0.64	64.00%
TOTAL	25	16	9	8	0.67	1	0.64	64.00%

Testing different OMR methods: SoundSlice

15 Broņislava Justa Rēzeknes Varaklānos, dzim. 1906; pier. E. Melngailis 1941. FS 1045, 9398.

1. Kur tu skrī-si, sy-la pē-lē? Rū-to, rū-tō!

1. Uz zēr-de ņu zēr-ņu zeg-tu, Rū-tō, rū-tō!

2. Kai tur teva nanūsissi?
— As ibēgšu aleņai.

3. Kai tu tur i nasasalsi?
— As sakuršu kaulu guni.

4. Kai tu tur i nasadegsi?
— Sasaspraužšu odoļosi.

Variant b of the song «Kur tu skrīsi sila pele» from the Emīls Melngailis book «Latviešu mūzikas folkloras materiāli, 2. sējums "Maliena" (1952)»

Kur tu skrī-si, sy-la pē-lē? Rū-to, rū-tō!

Uz zēr-de ņu zēr-ņu zeg-tu, Rū-tō, rū-tō!

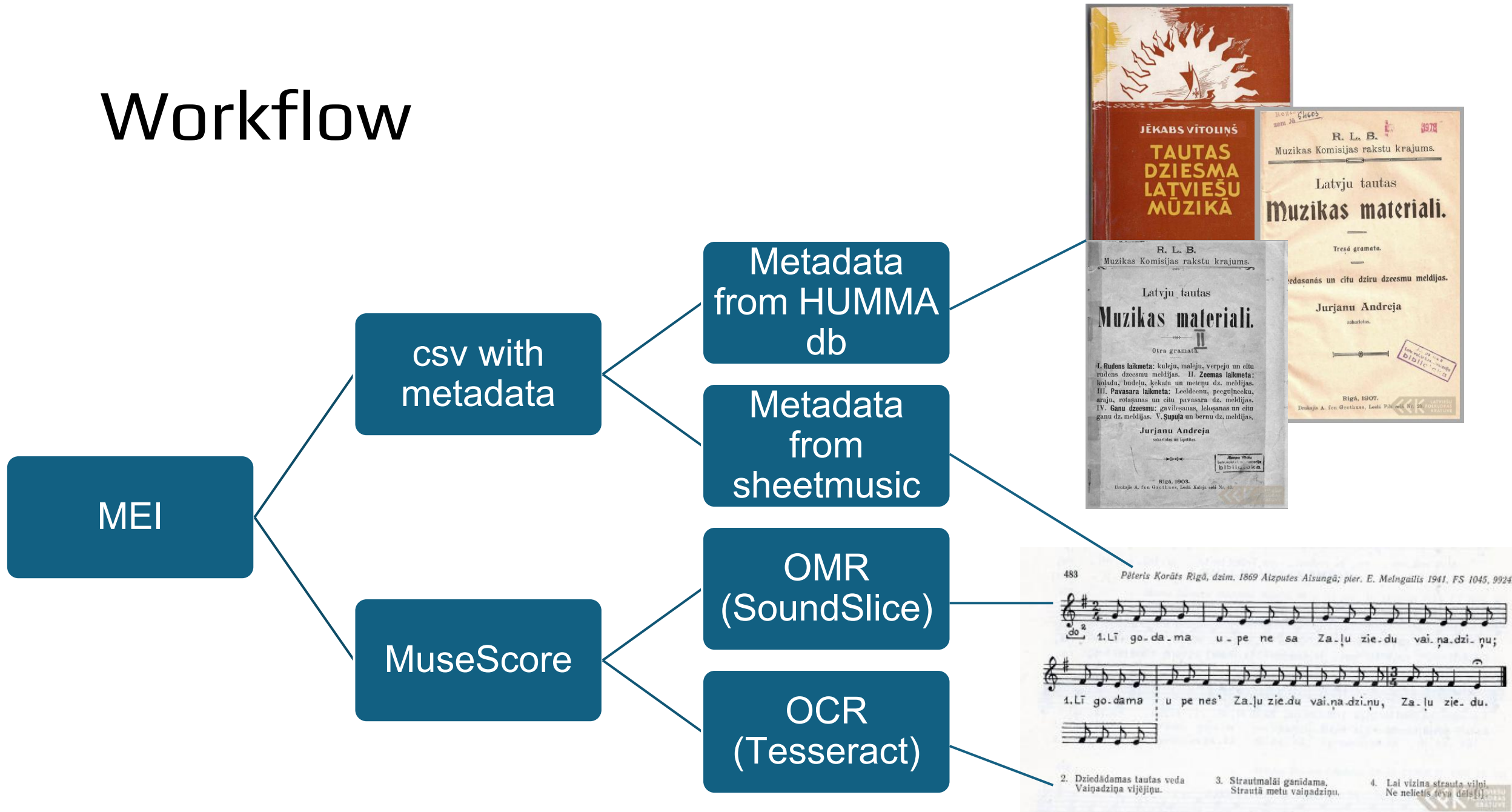
Variant b processed through SoundSlice

OMR								
Category	Total Items in the sample	True positives	False negatives (missing items)	False positives (extra items)	F1	Weights	Weighted accuracy	Percent
Pitch	24	24	0	0	1.00	0.40	0.40	40.00%
Rhythm	24	24	0	0	1.00	0.40	0.40	40.00%
Measures	8	8	0	0	1.00	0.13	0.13	13.33%
Key Signature	3	2	1	0	0.80	0.05	0.04	4.00%
Time Signature	1	1	0	0	1.00	0.02	0.02	1.67%
TOTAL	60	59	1	0	4.80	1.00	0.99	99.00%

OCR (neskaitot diakritiskās zīmes)								
Category	Total Items in the sample	True positives	False negatives (missing items)	False positives (extra items)	F1	Weights	Weighted accuracy	Percent
Metadata	1	1	0	0	1.00	0.04	0.04	4.00%
Syllables	24	15	9	9	0.63	0.96	0.60	60.00%
TOTAL	25	16	9	9	1.63	1.00	0.64	64.00%

OCR (skaitot diakritiskās zīmes)								
Category	Total Items in the sample	True positives	False negatives (missing items)	False positives (extra items)	F1	Weights	Weighted accuracy	Percent
Metadata	1	1	0	0	1.00	0.04	0.04	4.00%
Syllables	24	10	12	12	0.45	0.96	0.44	43.64%
TOTAL	25	11	12	12	1.45	1.00	0.48	47.64%

Workflow



How SoundSlice handles variants

The image shows a screenshot of the SoundSlice software interface. At the top, a modal dialog box titled "Should we include this?" is displayed. It contains two options: "Yes, include it" (selected with a blue radio button) and "No, ignore it" (unselected with a grey radio button). Below the dialog, a musical score is visible. A red dashed line with an arrow points from a specific musical phrase in the score to the "Yes, include it" button. The phrase in question is "Spožiem spangu", which is highlighted in a red box in the dialog. The main score consists of three staves of music in 2/4 time, with lyrics in Latvian. The first staff contains the lyrics: "Toai - sas ka - zas, tai - sas ka - zas, Nu bus ka - zas sSo - va - ka - re; Vi - Sa bra - E se - ta spi - de Nu Spo: Ziem vai - na - gle - mi,". The second staff contains: "sSa se - ta ka - zas, MNu bus ka zas to - Ya - ka - re, NU bUs le - zas. sSpi - de vai - na - gle - mi." The third staff contains: "Spo? iem pangu".

Should we include this?

Yes, include it ☒ No, ignore it ☐

s kā - zas, Nu būs kāzas šo - va - k
a spī - de vai - ņa - g

Spožiem spangu

Toai - sas ka - zas, tai - sas ka - zas, Nu bus ka - zas sSo - va - ka - re;
Vi - Sa bra - E se - ta spi - de Nu Spo: Ziem vai - na - gle - mi,

sSa se - ta ka - zas, MNu bus ka zas to - Ya - ka - re, NU bUs le - zas.
sSpi - de vai - na - gle - mi.

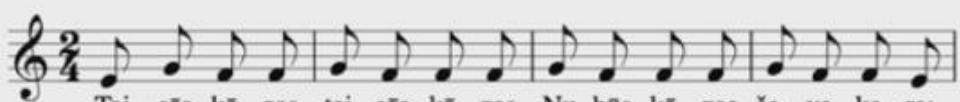
Spo? iem pangu

Final result in MEI

mei-friend Davis Engelis File Code View Manipulate Insert Help Taisas_kazas_V_II_1_41_A.mei

Taisās kāzas
Barba Grundmane Aizputes Alsungā, dzim. 1866 turpat; pier. E. Melngailis 1941.
FS 1045, 9956.

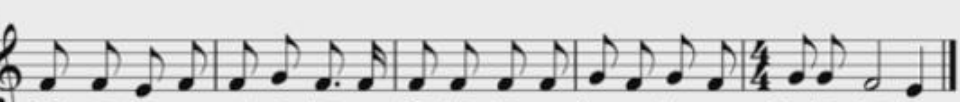
Voice



Tai - sās kā - zas, tai - sās kā - zas, Nu būs kā - zas šo - va - ka - re;

5

Vo.

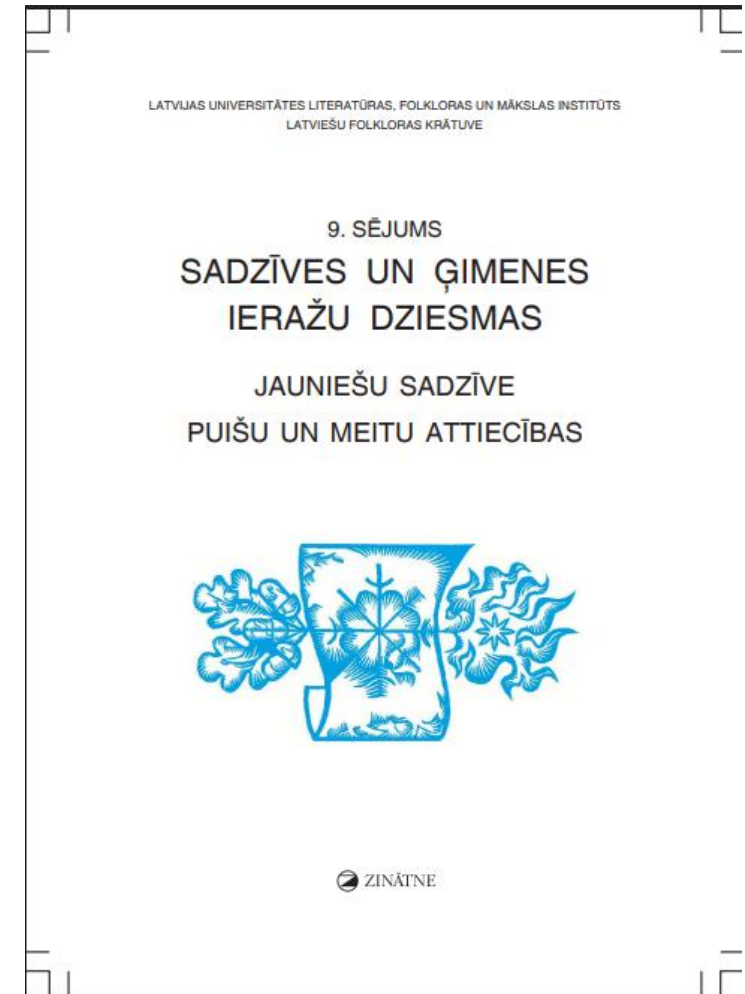


Tai - sās kā - zas, tai - sās kā - zas, Nu būs kā - zas šo - va - ka - re, Nu būs kā - zas.

```
1 <?xml version='1.0' encoding='UTF-8'?>
2 <xml-model href="https://music-encoding.org/schema/5.1/mei-basic.rng" type="application/xml" schematypens=
3 <xml-model href="https://music-encoding.org/schema/5.1/mei-basic.rng" type="application/xml" schematypens=
4 <?xml?><http://www.music-encoding.org/ns/mei" mei:version="5.1+basic">
5 <meiHead>
6 <fileDesc>
7 <titleStmnt>
8 <title type="main">Taisas kāzas, taisas kāzas</title>
9 <title type="subordinate">Subtitle</title>
10 <respStmnt>
11 <persName role="composer">Composer / arranger</persName>
12 </respStmnt>
13 </titleStmnt>
14 <pubStmnt>
15 <date isodate="2026-03-03T19:09:11"/>
16 </pubStmnt>
17 </fileDesc>
18 <otherMeta>
19 <meta name="textLanguage" content="latvian"/>
20 <meta name="witnessId" content="A"/>
21 <meta name="collector" content="Melngailis"/>
22 <meta name="PublishedID" content="#FMP 9, 1197"/>
23 <meta name="MelngailisSource" content="Latviešu mūzikas folkloras materiāli, 1. sējums &quot;Korsa&quot;
24 <meta name="VotlinsSource" content="V II , 1, 41"/>
25 <meta name="dateSource" content="1941"/>
26 <meta name="placeSource" content="Alsunga "/>
27 <meta name="placeSourcePast" content="Aizputes Alsungā"/>
28 <meta name="SingerSurname" content="Grundmane"/>
29 <meta name="SingerName" content="Barba"/>
30 <meta name="SingerBirthYear" content="1866"/>
31 <meta name="SingerBirthPlace" content="Alsunga "/>
32 <meta name="Encoder" content="Dāvis Engelis"/>
33 <meta name="OmrTool" content="Soundslice"/>
34 </otherMeta>
35 </meiHead>
36 <music>
37 <body>
38 <mdiv>
39 <score>
40 <scoreDef>
41 <pgHead func="first">
42 <rend haligh="center" valign="top">
43 <rend type="title" fontsize="x-large">Taisās kāzas</rend>
44 <lb/>
45 <rend type="subtitle" fontsize="large">Barba Grundmane Aizputes Alsungā, dzim. 1866 turpat;
46 </rend>
47 </pgHead>
48 <staffGrp>
49 <staffDef n="1" lines="5" meter.count="2" meter.unit="4">
50 <label>Voice</label>
51 <labelAbbr>Vo.</labelAbbr>
52 <instrDef midi.instrnum="52"/>
53 <clef shape="G" line="2"/>
```

Textual variation recovery (using LLM)

- The dataset contains ~45 000 samples of text.
- The text variants are written in an abbreviated form.
- The challenge: automate text variant recovery in full form.



Text variant reconstruction using Gemini

We tested several LLMs on the task of generating expanded-format transcriptions from spoken song inputs. The models varied significantly in how they handled constraints and how often they introduced hallucinated content.

Vīgli mani žogi rōve,
Tōli mani pīminēja:
Pīmiņ mārša, molt īdama,
Mōte – vītu taiseidama. L 548, 13939
Z≈2, V≈5, L≈7

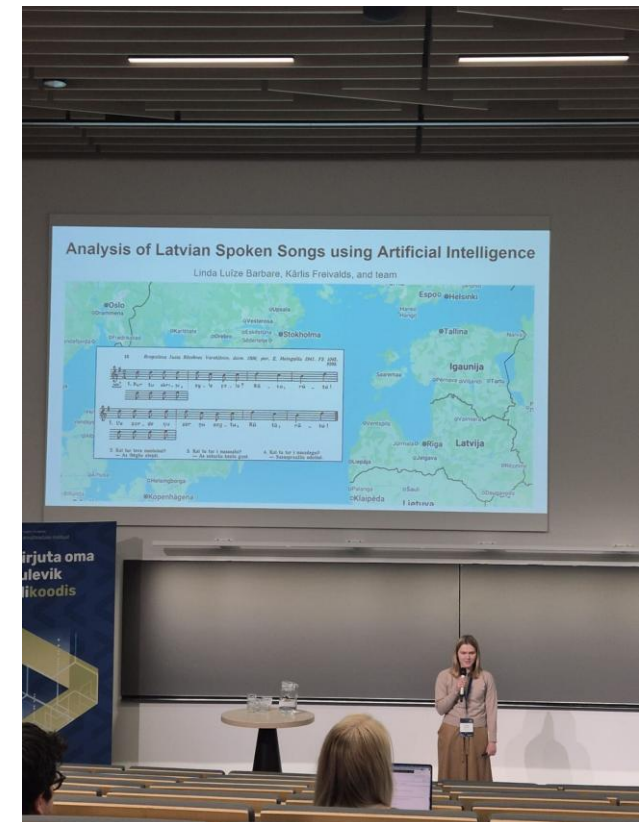
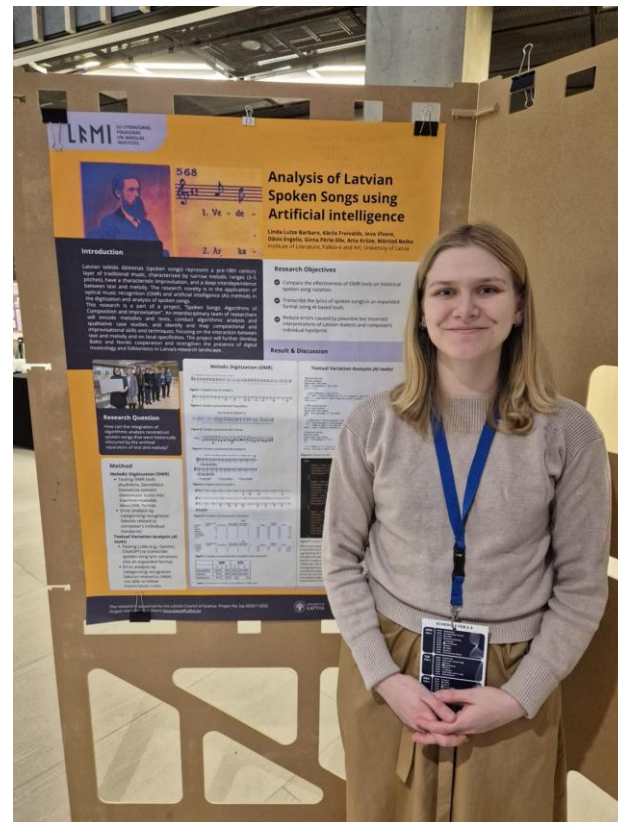
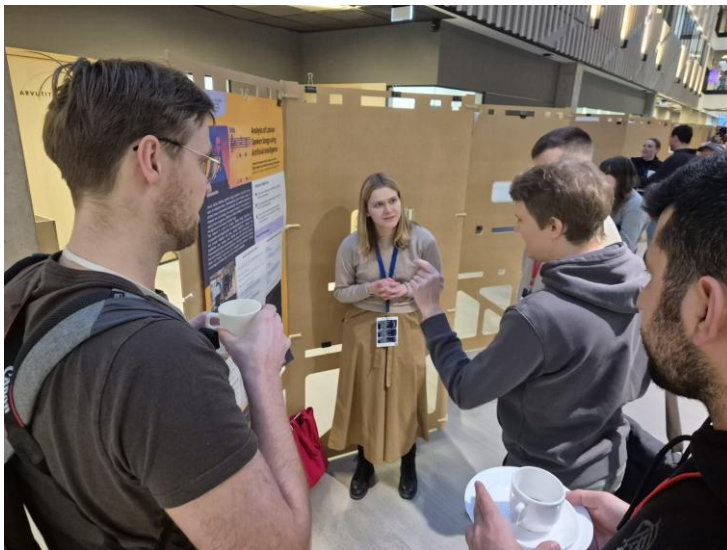
4. putru vōreidama L 1422, 69, N≈1.
L: 4. Mōt' – azaidu vōrīdama 494, 711, ≈3; 3. moltyvē 4. putru vōrīdama 509, 343; 4. gultu 828, 18219, ≈1; 1. žogs nūrōve 3. mōte 4. Pīmiņ 935, 23747; 3. moltuvē 1142, 682, ≈3; 4. ustabu slaucēdama 1249, 754, ≈4.

Gemini produced the most promising results.
It's outputs aligned 100% with the original material.

Dziesmas_Nr	Tips	References_Nr	Dziesmas_teksts	Kulturvesturiskai	Kolekcijas_Nr	Vienibas_Nr	Pielidzinatas_dz	Uncertainty_Flaç	Confidence_Sco	Kludas_jemesls	Variācijas_kodej	PD_skaits
58309	Pamatdziesma	0	Es bij puika, man bij nauda, Es ar naudu lielījos: Kad uz krogu dancot gāju, Naudu bērū zābakā, Kad piesītu es kājiņu, Tad naudiņa nožvakšķēja.	Z	185	28		FALSE	100			0
58310	Pamatdziesma	0	Es bij puika, man bij nauda, Es varēju lielīties; Es varēju meitas viit No krodziņa uz krodziņu.	L	1950	3040	L≈1	FALSE	1			1
58311	Pamatdziesma	0	Šītie tādi mulķa zēni, Zin tik dzert un lielīties; Ne tiem jēgas, ne prātiņa, Ne ar' laba tikumiņa.	K	Bb 20	887	V≈1, N≈1	FALSE	1			2
58311	Pamatdziesma	0	Šītie tādi mulķa zēni, Zin tik dzert un lielīties; Ne tiem jēgas, ne prātiņa, Ne ar' laba tikumiņa.	K	Bb 20	887	V≈1, N≈1	FALSE	1			2
58311	Variācija	0	Šītie tādi mulķa zēni, [Māk] tik dzert un lielīties; Ne tiem jēgas, [ne tiem prāta], [Ne labā] tikumiņa.	V	72	2303		FALSE	0.9		2. Māk 3. ne tien	0
58312	Pamatdziesma	0	Kas tie tādi tēva dēli, Krogā māka lielīties? Būt' sējuši, būt' plāvuši, Būt' zinājši tikumiņa.	Z	929	67006		FALSE	1			0

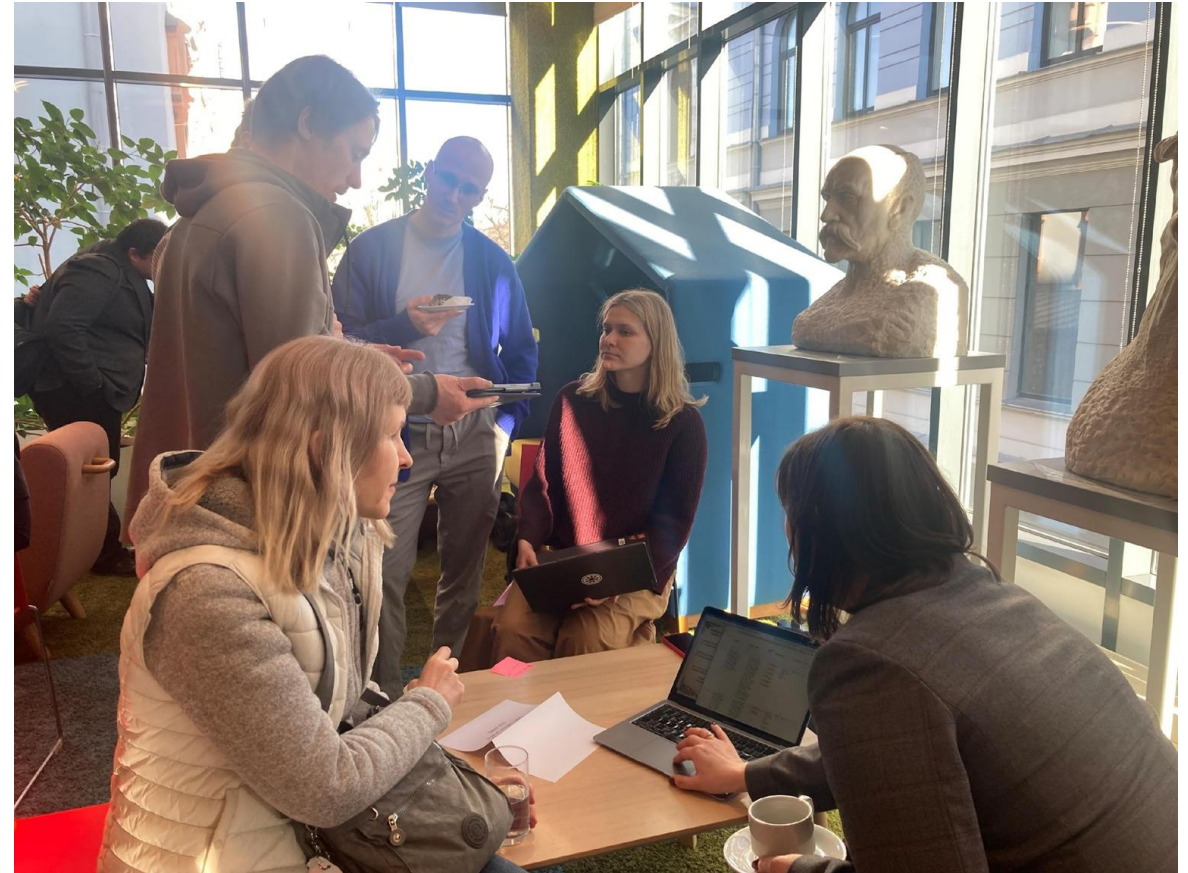
Participation in conferences and workshops

Presented poster in the international Workshop on Uncertainty in Machine Learning (WUML 2026), held at the University of Tartu from February 2 to 4.



Participation in conferences and workshops

Demonstrated solutions for the digitization and analysis of Latvian folk songs in CLARIN–DARIAH Latvian Conference Highlights Developments in Digital Humanities Infrastructures



Future work

1. Write a paper about the text variant reconstruction (paper hand in til 30.06.2026) for EuroMed2026, the 11th International Conference on Digital Heritage, in Limassol, Cyprus, 2. - 5. November 2026. Accepted papers will be posted on Springer
2. Think about how to increase the novelty of this work.