

LMI LU LITERĀTURAS
FOLKLORAS
UN MĀKSLAS
INSTITŪTS



Kā dators iemācījās izgriezt dziesmas no nošu grāmatām

Lai iekodētu teiktās melodijas, vispirms bija veicams pavisam praktisks uzdevums: no nošu izdevumiem bija jāizgūst vajadzīgās melodiju vienības. Pagaidām strādājām ar Jēkaba Vītoliņa sastādīto sēriju "Latviešu tautas mūzika". Ar roku tas būtu apmēram tūkstošiem rūpīgi izgriežamu attēlu, tāpēc nolēmām šo darbu automatizēt. Kāpēc tas nav elementārs uzdevums?

Grāmatas ir skenētas kā attēli bez teksta slāņa, tāpēc vispirms bija nepieciešama teksta atpazīšana (OCR). Un tieši te sākas grūtības, jo: 1) nošu lapa OCR programmai ir haoss, notis rada tūkstošiem viltus "burtu"; 2) vienību robežas datoram bez paskaidrošanas nav saprotamas; 3) katrs sējums maketa ziņā ir citāds. Risinājums bija Python skriptu komplekts, kas darbojas trīs soļos. Vispirms katra grāmatas lapa tiek nolasīta ar atvērtā koda Tesseract OCR latviešu valodas modeli, saglabājot katra atpazītā vārda koordinātas. Pēc tam skripts katrai datubāzes dziesmai meklē atbilstošo numuru grāmatas lapas kreisajā malā ar nosacījumu, ka šim numuram blakus atrodas kaut kas metadatiem līdzīgs, lai nesajauktu melodijas numuru, piemēram, ar panta numuru. Grūtākais datoram ir saprast, kur dziesma beidzas, tāpēc bija jāapraksta, kā vizuāli atpazīt jebkuru maketa elementu un nesajaukt to ar citiem. Trešais solis ir metadatu izvilkšana. Šaurā metadatu josla tiek nolasīta vēlreiz ar augstāku izšķirtspēju, vairāki nolasījumi tiek attīrīti un salīdzināti ar datubāzē jau iekļauto informāciju, nonākot pie precīzākas informācijas. Nolasītie metadati pēc tam tika pārbaudīti manuāli – cilvēka veikta validācija ir svarīga jebkurā solī. Secinājām, ka biežākās OCR kļūdas rodas vietās, kur burtzīmes ir vizuāli līdzīgas, piemēram, slīpsvītra tiek sajaukta ar burtu "l" vai burts "f" tiek sajaukts ar burtu "j".

Galarezultāts bija iekodēšanai izgrieztas dziesmas un datubāzē papildināti metadati. Interesantākais šajā darbā nebija OCR, bet gan tas, cik daudz paskaidrojumu vajag, lai datoram pateiktu, kur beidzas viena dziesma un sākas otra. Cilvēks to redz uzreiz, bet datoram tas jāizskaidro ar pārdesmit likumiem. Toties, kad likumi bija uzrakstīti vienam sējumam, pārējie pieci prasīja tikai dažas rindiņas pielāgošanas. Tūkstoš attēlu darbs, kas prasītu vairākas dienas manuāla darba, tika paveikts dažās stundās, un skripts ir gatavs nākamajam sējumam, ja tāds parādīsies.

Detalizētāku un tehniskāku aprakstu par šo procesu un rezultātiem varat lasīt [šeit](#).

Kārlis Freivalds

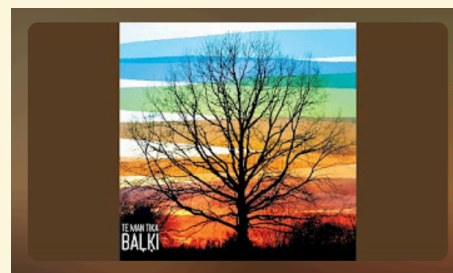
Ko tā blusa augstu lec?! –
"Zinātnieku nakts"
LU Botāniskajā dārzā!



Šogad mūs un tautasdziesmām veltīto LU LFMI "Folk-loras stūrīti" meklējiet 25. septembrī LU Botāniskajā dārzā, Rīgā, Šteinhauera ielā 2, kur kopā ar Latvijas Universitātes biologiēm pievērsīsimies parazītu pasaulei: kopā centīsimies blusas, utis u. c. parazītus rādīt kā veselumu – gan reālas dzīvības būtnes, gan arī kā daļu no tautasdziesmu poētiskās pasaules. Gaidīsim Jūs!

Iesakām!

Rotaļa "Dindaru, dandaru ozoliņi" 15 radošās interpretācijās tautas, akadēmiskās un populārās mūzikas žanros:



(etnomuzikoloģes Ievas Vīveres YouTube kanāls)

Saite:

<https://www.youtube.com/playlist?list=PLf0xxNfV805E>

Nošu korpusa veidošana

Mēs virzāties tālāk ar teikto melodiju iekodēšanu. Vispirms maza atkāpe par to, ko tas nozīmē. Ņem nošulapu, iebaro to algoritmam un ieraugi visas notis un pavadinformāciju datora ekrānā, rediģējamā formātā. Nošulapa, savā pirmējā veidolā viegli izlasāma muzikāli trenēta cilvēka acij, kļūst mašīnlasāma. Angliski tam parocīgs akronīms: OMR – *optical music recognition* jeb optiskā mūzikas atpazīšana. Pastāv jau gatavi OMR rīki, ar kuriem nošu iekodēšana kļūst ievērojami ātrāka, nekā tad, ja pāri tūkstotim teikto dziesmu mēs iekodētu savām rokām noti pa notij, zilbi pa zilbei. Bet ik nošu izdevums met ceļā savus izaicinājumus. Aiz katra nošu izdevuma ir savi tipogrāfiskie risinājumi, lapas izkārtojums, nošzīmes un burtveidoli, zīmūlpietīmes un laika zoba pēdas, arī ieskenētā attēla kvalitāte. Jēkaba Vītoliņa sastādītie “Latviešu tautas mūzikas” sējumi, mūsu nošu korpusa galvenais avots, vēl nāk ar dažādiem dziesmu numerācijas risinājumiem un nekonsekventu, daudzveidīgu dziesmu variantu pierakstu. No OMR skatpunkta – arī starptautiskā mērogā – melodisko variantu pieraksti Vītoliņa sējumos ir unikāli. Tā ir jauna problēma, un mēs esam pirmie, kas meklē tai risinājumus.

Dažas mūsu uzkrātās atziņas: no OMR rīkiem, programmām ir tādas, kas Vītoliņa grāmatas iekodē ar apbrīnojamu precizitāti, ja skatāmies nošu augstumus. Tiktāl 100% precizitāte. (Galvenokārt šī iemesla dēļ iekodēšanai lietojam maksas programmu *Soundslice*.) Gandrīz tikpat labi ir ar nošu ilgumiem, aptuveni 98% iekodēto nošu ir precīzi atveidotas. Ne tik labi veicas ar artikulācijas simboliem, bet arī te rādītāji parasti ir iepriecinoši, mums maz darba, šādas kļūdas labojot. Lielākā sāpe par to, ka patlaban jau gatavie, izstrādātie OMR rīki neatbalsta latviešu valodu, tādēļ visas diakritiskās zīmes līdzās citiem programmas kļūdaiņiem atpazītiem burtiem mums jāizlabo manuāli. Bet arī te nedaudz palīdz algoritmi, mēs iedarbinām atsevišķu kodu (rakstzīmju atpazīšanu veic *Tesseract* OCR pakotne), kas latviešu alfabētu atpazīst itin veikli, un atkal – mums paliek mazāk kļūdu labošanas darba. Ja lasītājam interesē, ar kādiem kodiem strādājam, sazinieties ar mums, iedosim saiti uz GitHub repozitoriju.



The screenshot displays the 'mei-friend' web application interface. The top navigation bar includes links for Login, File, Code, View, Manipulate, Insert, and Help. The main content area is split into two panels. The left panel shows a musical score for 'Track 1' in 2/4 time, with lyrics in Latvian: 'Vis - ta cā - līt' - at - ūkt - ru - si Ga - ra - jā - si ka - pe - pēs.' The right panel shows the corresponding XML code, which includes metadata, revision information, and musical notation details. The bottom panel shows a zoomed-in view of the musical notation and lyrics, with a list of lyrics below the staff.

Nošu iekodēšanas paraugs – digitalizētais nošu piemērs, tā kods (pa labi) un iekodētais avots

leguvumi no visa šī darba – papīra nošu korpus iegūst savu digitālo versiju. Mūsu uzdevums to izveidot visiem pieejamu. Tāpat varēsīm ņemt talkā datoradarbināmus mūzikas analīzes rīkus un strādāt ar lielu datu apjomu – visām līdz šim pierakstītajām teiktajām melodijām. Līdz tam vēl daudz iekodēšanas darba. Šobrīd esam iekodējuši apmēram trešdaļu no teikto dziesmu nošu korpusa. Svarīgākais secinājums – datu sagatavošanai, kas šajā gadījumā ir digitālā nošu korpusa izveide, projekts ir jāparedz daudz vairāk laika! To, ka datu sagatavošana (digitalizēšana, tīrīšana u. tml.) digitālos projektos ir laikietilpīgākais uzdevums, apliecinās arī citi digitālo humanitāro zinātņu projektu īstenotāji.

Tradicionālās dziedāšanas skola Drabešos



Jau daudzus gadus vasarās Drabešos pulcējas tie, kurus aizrāvusi tradicionālā dziedāšana, lai kopā apgūtu un papildinātu gan praktiskās iemaņas, gan arī teorētiskās zināšanas šajā jomā. Šogad šāda tikšanās bija 7.–9. augustā, un tajā bija iespēja piedalīties arī mūsu projekta pētniekiem. Savos priekšlasījumos Ieva Vīvere un Ginta Pērle-Sīle informēja par projektu, jau paveikto. Tomēr lielākā priekšlasījumu daļa bija veltīta pazīstamajai rotaļai "Dindaru, dandaru ozoliņi", to apskatot no mūzikas un teksta variativitātes viedokļa, iezīmējot sakarības teksta un melodijas "draudzībā".

Ginta Pērle-Sīle un Ieva Vīvere

Zināma DR Kurzemes melodiskā tipa detektīvs

Lai arī tas plānots tikai projekta trešajā gadā, jau tagad nepacietīgi ķērāties pie jautājuma par teikto dziesmu reģionālajiem tipiem un īpatnībām. 2003. gada grāmatā “Latviešu teikto dziesmu melodiskā veidojuma tehnikas” Mārtiņš Boiko iekļāva analītiskas etīdes par atsevišķu novadu un pagastu teikto melodiju tipiem. Šīs etīdes nu sākam attīstīt un padziļināt, pārnesot interaktīvā digitālā vidē, precīzāk, kartējot un salīdzinot areālus, melodiju analīzi papildinot ar dziesmu tekstu izpēti un avotu kritiku. Šī topošā metode ved pie jaunu kopsakarību saskatīšanas un jaunām aizraujošām mīklām.

Pateicoties aicinājumam uz Drabešu tradicionālās dziedāšanas skolu, detalizētāk apskatījām melodisko tipu, kura areāla centrs atrodas uz ziemeļiem no Liepājas ([te areāls aplūkojams kartē](#)). Šis melodiskais tips tautā pazīstams, pateicoties rotaļai “Dindaru, dandaru ozoliņi”. Boiko apzīmējumu lietojot, tas ir DR Kurzemes *divzonu meloss*. Tam raksturīga dziedāšana pamīšus divās augstuma zonās – pirmajā panta pusē augstāk un otrajā panta pusē melodijai nokrītot zemāk. Zināmajā rotaļā pierasts šo melodiju dziedāt svinīgi un līdzeni, noteiktā skaņu augstumā vispirms augstāk un pēc tam kvartu zemāk. Taču visu avotu salīdzināšana rāda, ka tradicionāli dziedāšana visticamāk bijusi daudz dinamiskāka un nav fiksējama konkrētos skaņu augstumos. Izteikti “runātīgi” un notīs teju nepierakstāmi 1961. gada ieskaņojumos dzied [Grieta Bekere no Kapsēdes](#). Kopā par šo melodisko tipu liecina ap 25 nošu pieraksti un ap 19 minūšu gari skaņu ieraksti no Kapsēdes un Vērgales. Tie parāda lielu dažādību melodisko intervālu ziņā un, iespējams, grūtības šo dinamisko dziedāšanu pierakstīt notīs. Turklāt avoti liecina arī par garu vilkšanu, akcentiem un izsaucieniem panta beigās. Visu avotu apkopojums šo lokālo tradīciju rosina uztvert kā melodisku modeli vai melodisku ideju, nevis konkrētu melodiju, un sakārtotas “variēšanas” vietā rosina domāt par “aptuveno dziedāšanu” šī tradicionālā modeļa ietvarā, atceroties jakutu izcelsmes etnomuzikologa Eduarda Aļeksejeva (1937–2021) “zonālās intonēšanas” jēdzienu.

Šī melodiskā korpusa tekstu kartēšana un izpēte ļauj padziļināt izpratni par tradīciju. Šo melodisko tipu pārsvarā dziedāja godos, dinamiskās, bieži ar kustību saistītās norisēs – pādes dīdīšanā un apdziedoties ķekatās, kāzās. Ar šo melodiju nav fiksēti teksti no sadzīves ieražām jeb ikdienas norisēm. Par teiktajām melodijām etnomuzikologi domā kā par daudzfunkcionālām, proti, ka vienu melodiju varēja izmantot dažādās situācijās. Šis materiāls liek domāt, ka uz šo patiesību vērts skatīties niansētāk. Varbūt viens melodiskais tips ir kalpojis vairākām, taču ne visām situācijām? Tāpēc pētījuma nākošais solis būs salīdzināt šo divzonu melodisko tipu kopsakarībā ar citām teiktajām melodijām šajā areālā, pievēršot uzmanību dziesmu tekstiem.

“

Teksta fragmentu kartēšana padziļina izpratni arī par to, ko varam uzskatīt par “tradīcijas areālu”. Melodiskais areāls šajā gadījumā ir šaurs; tekstu virknējumi – unikāli katrai situācijai. Taču īsāki tekstu motīvi var būt plaši izplatīti visā Latvijā, gan ar lokālām variācijām. Piemēram, lai gan dažādas darbības pādes likteņa ietekmēšanai (formula “lai manai pādītei”) ir sastopamas visā Kurzemē un arī Zemgalē, tomēr dejošana pa “pusplānu, lai mana pādīte pusgadu staigā” pamatā [sakrīt ar divzonu melosa areālu](#). Turklāt šāda ģeogrāfija ir gan “Latvju dainās” iekļautajiem tekstiem (11 teksti), gan arī vēlāk savāktajiem, kas atrodami akadēmiskajā izdevumā “Latviešu tautasdziesmas” (18 teksti). Savukārt pārsteigums bija, ka kažoka vilkšana, “lai manai pādītei sprogaini mati”, izteikti koncentrējas Vidzemes vidienē. Varam tikai minēt, kāpēc.

Arvien lielāku uzmanību pētījumā veltām avotu kritikai jeb jautājumiem, kurš, kāpēc un kādu materiālu ir pierakstījis vai publicējis. Mūsu uzmanību piesaistīja tas, ka kanonizējusies rotaļa “Dindaru, dandaru ozoliņi” ir salikta no diviem pierakstiem, kas publicēti Jurjānu Andreja “Latvju tautas mūzikas materiālu” 4. grāmatā 1912. gadā. Šo salikumu kopš 20. gs. 80. gadu sākuma popularizēja folkloras kustības dalībnieki. Pazīstamā teksta analīze liek domāt, ka tas varēja būt kaut kādā mērā jaunradīts. Pagaidām mums nav izdevies tikt līdz šīs rotaļas “personas kodam”, taču atradām, ka jau pirms Jurjānu Andreja šī

rotaļa ar nosaukumu “Ozoliņš” 1893. gadā publicēta Paleviču Indriķa “Dziedāšanas mācībā tautas skolām”. Varbūt vienkāršais un statiskais melodiskais variants izvēlēts pedagoģiskos nolūkos kā vieglāk apgūstams?

lewa



Rotaļa “Ozoliņš” Paleviču Indriķa “Dziedāšanas mācībā tautas skolām” (fragments)

Mums vēl nav atbilžu uz visiem jautājumiem, taču jau tagad esam gatavas par šo visu stāstīt detalizētāk un ar piemēriem. Aiciniet mūs!

Tautasdziesmām un tradicionālajai dziedāšanai veltīta konference Tartu

Pašā augusta izskaņā, no 24. līdz 27. augustam, projekta dalībnieces Arta Krūze un Ginta Pērle-Sīle piedalījās Igaunų Folkloras krātuves organizētajā 54. starptautiskās balāžu komisijas un 16. tradicionālajai dziedāšanai veltītajā konferencē "Dziesmas arhīvos un kopienās: teksti un dzīvā tradīcija".

Konferences dalībnieki sprieda par jaunākajiem atklājumiem tautasdziesmu pētniecībā, dalījās ar labās prakses piemēriem tradicionālās dziedāšanas mācībās. Nozīmīga vieta konferencē bija ierādīta tēmai par dziesmām un dziedāšanu kā reakciju uz krīzes situācijām un izaicinājumiem, ko šādas krīzes rada nemateriālā kultūras mantojuma saglabāšanai. Neiztrūka arī pārdomas par jauno tehnoloģiju radītajām iespējām un izaicinājumiem dziesmu folkloras pētniecībā.

LU LFMI pārstāves konferencē piedalījās ar referātu "Divrinžu kā strukturālu vienību atkārtotāšanās latviešu tautasdziesmās: ieskats lielapjoma tekstu korpusā" (*Recurring Couplets as Structural Units in Latvian Folk Songs: Insights from a Large-Scale Textual Corpus*), stāstot par rezultātiem, kas iegūti, tautasdziesmu analīzē izmantojot automatizētu tekstu salīdzināšanas rīku.

Ar konferences programmu un referātu kopsavilkumiem iespējams iepazīties šeit: <https://www.folklore.ee/KfV2026/>



Konferences dalībnieku kopbilde. Foto: Enn Kalev Tarto

Dāvis Enģelis stāsta par teiktajām dziesmām "WoRMS" darbnīcā Vīnē

LU LFMI zinātniskais asistents Dāvis Enģelis piedalījās 7. Starptautiskajā mūzikas sistēmu nola-sīšanas darbnīcā "WoRMS" (7th International Workshop on Reading Music Systems), kas šogad norisinājās Vīnē Starptautiskās dokumentu analīzes un atpazīšanas konferences (ICDAR) ietvaros. "WoRMS" ir viens no nozīmīgākajiem notikumiem pētniecības jomā, kas pievēršas nošu raksta digitālai nolasīšanai un iekodēšanai – algoritmu spējai atpazīt un interpretēt muzikālo pie-rakstu. Dāvis Enģelis darbnīcā uzstājās ar referātu "Interpretējot un iekodējot melodiskos variantus: Latvijas tradicionālo rečita-tīvu gadījums" (*Interpreting and Encoding the Melodic Variability: The Case of Latvian Traditional Recitatives*), pievēršoties Latvijas tradicionālās mūzikas materiāla interpretācijas un digitālās iekodēšanas jautājumiem.

[Vairāk par darbnīcu ŠEIT.](#)



Dāvis Enģelis konferencē Vīnē

Projekts "Teiktās dziesmas: kompozīcijas un improvizācijas algoritmi" (projekta nr. Nr. Izp-2025/1-0252)

SAZIŅAI: ieva.vivere@lulfmi.lv |

[PAR PROJEKTU](#) |

[VĒSTKOPAS ABONĒŠANA](#)