From Corpus to Classroom: Interactive Access to Children's Song Repertoire on the MUSCO Platform

Vanessa Nina Borsan*, Faculty of Arts, University of Ljubljana,
vanessanina.borsan@ff.uni-lj.si
Jure Juvan, Faculty of Computer and Information Science, University of Ljubljana,
jure.juvan@fri.uni-lj.si
Matija Marolt, Faculty of Computer and Information Science, University of Ljubljana,
matija.marolt@lgm.fri.uni-lj.si
Matevž Pesek, Faculty of Computer and Information Science, University of Ljubljana,
matevz.pesek@fri.uni-lj.si
Leon Stefanija, Faculty of Arts, University of Ljubljana,
leon.stefanija@ff.uni-lj.si
Filip Trplan, Faculty of Computer and Information Science, University of Ljubljana,
filip@trplan.si

Abstract

Singing is often one of the first ways children engage with music. Yet, the digital tools supporting this experience rarely efficiently satisfy the pedagogical, linguistic, or cultural contexts in which songs are taught. Most digitised music corpora and their search systems rely on generalised models of pattern analysis, offering little support for selecting repertoire suited to specific age groups, vocal ranges, or native-language environments. This contribution proposes a methodology for filtering music corpora, grounded in the intended use of the repertoire. We introduce eight pedagogical filters—four melodic (based on ambitus and intervallic motion) and four rhythmic (focused on metre, breaks, and note values)—each implemented across at least two levels of complexity and informed by national education curricula and pedagogical research. These filters enable age-appropriate (from years 6 to approximately 10), context-sensitive access to music, aligning selection with children's vocal and cognitive development. To demonstrate this approach, we have curated a new open corpus of 123 Slovenian children's songs from *Ciciban*, a long-running educational magazine for children. The corpus is encoded and integrated into the emerging MUSCO platform, where the filters are embedded in an interactive interface that supports both educators and researchers, regardless of technical or pedagogical background. Evaluation using classification metrics across children's and folk song corpora shows that the filters consistently exclude more songs from the non-children's corpus, making a strong case for their pedagogical validity. Finally, the platform and its respective corpora also enable a variety of pattern-based music analysis (metadata, intervallic or rhythmic and other queries), through suffix-array-based, n-gram or other search structures. All code, datasets, and the evaluation notebook are openly released. In addition to supporting early music education, our methodology and tools enable context-aware corpus design for pedagogical, musicological, and computational research.

Selected References

Betty W. Atterbury. 1984. Children's Singing Voices: A Review of Selected Research. B*ulletin of the Council for Research in Music Education* 1, 80 (1984), 51–63.

Charles Ballester, Baptiste Bacot, Louis Bigo, …, and Mathieu Giraud. 2025. Interacting with Annotated and Synchronized Music Corpora on the Dezrann Web Platform. *Transactions of the International Society for Music Information Retrieval* 8, 1 (2025), 121–139.

Bogdana Borota and Alenka Kovačič Divjak. 2015. *Dejavnosti glasbenega opismenjevanja.* Univerzitetna založba Annales, Koper.

Vanessa Nina Borsan. 2024. *Melodic and Descriptor Patterns: A Computational Approach for Digitisation, Annotation, and Analysis of Slovenian Folk Song Ballads.* Ph.D. Thesis. Université de Lille.

Nádia Carvalho, Sara González-Gutiérrez, Javier Merchán Sánchez-Jara, G. Bernardes, and M. Navarro-Cáceres. 2021. Encoding, Analysing and Modeling I-Folk: A New Database of Iberian Folk Music. *Proceedings of the 8th International Conference on Digital Libraries for Musicology* (DLfM'21), 75–83.

Soonbeom Choi, Wonil Kim, Saebyul Park, Sangeon Yong, and Juhan Nam. 2020. Children's Song Dataset for Singing Voice Research. In *Extended Abstracts for the Late-Breaking Demo Session of the 21st International Society for Music Information Retrieval Conference*, Vol. 4. ISMIR, Montréal, Canada, 1–2.

Olga Denac. 2002. *Glasba pri celostnem razvoju otrokove osebnosti: priročnik za vzgojitelje, razredne učitelje, učitelje glasbe in glasbenih predmetov v splošnih in glasbenih šolah*. Zavod Republike Slovenije za šolstvo, Ljubljana.

Behnam Faghih and Joseph Timoney. 2022. Annotated-VocalSet: A Singing Voice Dataset. *Applied Sciences* 12, 18 (2022), 9257.

Vassilis Katsouros, Evita Fotinea, Renaat Frans, …, and Marcus Liwicki. 2018. *Designing for the User Experience in Learning Systems*. Springer, Switzerland, Chapter iMuSciCA: Interactive Music Science Collaborative Activities for STEAM Learning, 123–154.

Ivan Lešnik. 2015. Children's Singing in Slovenia: Vocal Range Analyses. In *Contemporary Approaches to Music Teaching and Learning*. LAP LAMBERT Academic Publishing, Saarbrücken, Lambert, 127–141.

Tania Caamaño Liñares, Jesús Rodríguez Rodríguez, Montserrat Castro Rodríguez, and Diana Marín Suelves. 2023. Digital didactic resources and music: mapping the last decade of research. *Music Education Research* 25, 4 (2023), 351–366.

Anna Wood, Kathryn R Kirby, Carol Ember, Stella Silbert, Sam Passmore, Hideo Daikoku, John McBride, Forrestine Paulay, Michael Flory, John Szinger, et al. 2021. The Global Jukebox: A public database of performing arts and culture. *PLOS One* 17, 11 (2021), e0275469.

Ji-Hyun Yi. 2022. A Review of Literature on the Music-using Apps and Developmental Strategies. *Childhood Education and Culture* 26, 4 (2022), 403–428.

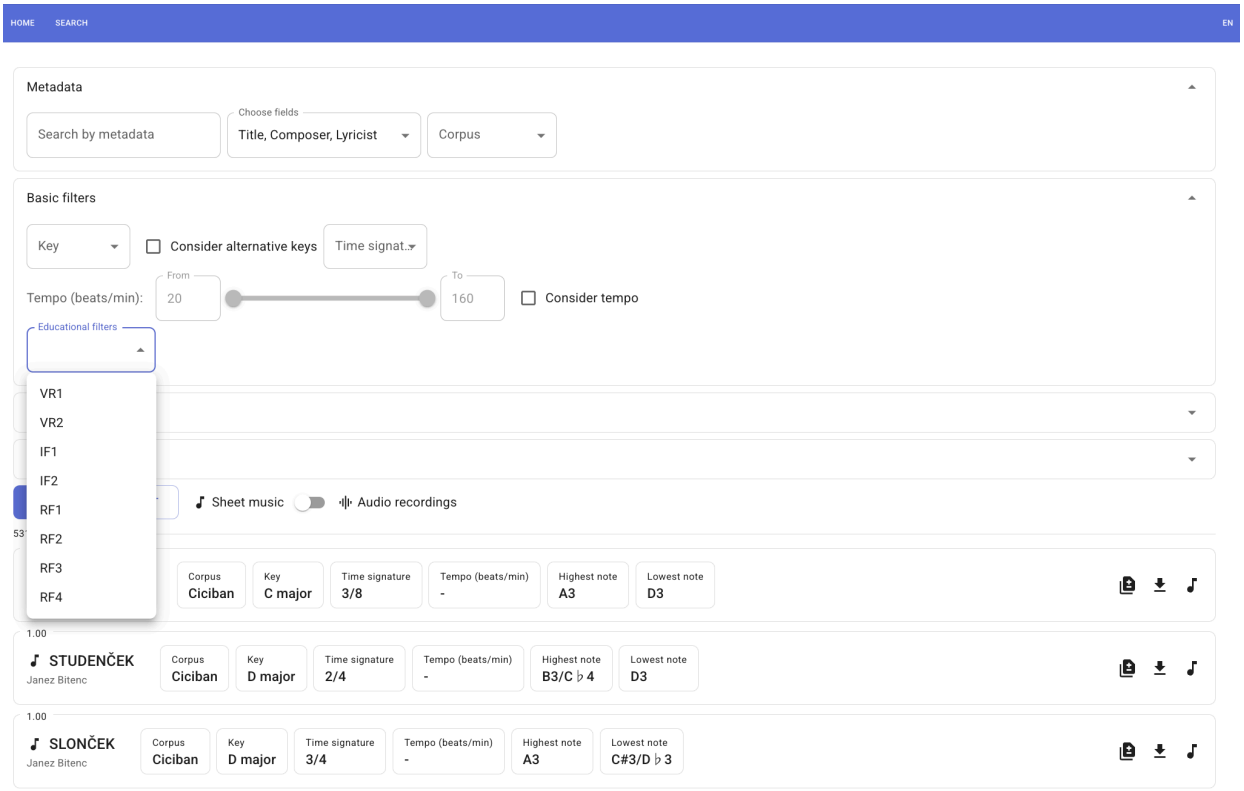

Figure 1. *Interactive filtering interface of the MUSCO platform.* The MUSCO platform allows users to search and filter children's songs based on metadata, musical features (key, tempo, time signature), and educational filters aligned with developmental stages. The results shown in the figure include several songs from the *Ciciban* corpus, along with their metadata, options to download sheet music and metadata, and the ability to visualise the score directly on the platform by clicking on an individual song title.