View Reviews

Paper ID

32

Paper Title

Generating Symbolic Music from Natural Language Prompts using an LLM-Enhanced Dataset

Track Name

Papers

Reviewer #1

Questions

2. I am an expert on the topic of the paper.

Agree

3. The title and abstract reflect the content of the paper.

Agree

4. The paper discusses, cites and compares with all relevant related work

Agree

5. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chosen, otherwise write "n/a")

n/a

6. Readability and paper organization: The writing and language are clear and structured in a logical manner.

Agree

7. The paper adheres to ISMIR 2025 submission guidelines (uses the ISMIR 2025 template, has at most 6 pages of technical content followed by "n" pages of references or ethical considerations, references are well formatted). If you selected "No", please explain the issue in your comments.

Yes

8. Relevance of the topic to ISMIR: The topic of the paper is relevant to the ISMIR community. Note that submissions of novel music-related topics, tasks, and applications are highly encouraged. If you think that the paper has merit but does not exactly match the topics of ISMIR, please do not simply reject the paper but instead communicate this to the Program Committee Chairs. Please do not penalize the paper when the proposed method can also be applied to non-music domains if it is shown to be useful in music domains.

Strongly agree

9. Scholarly/scientific quality: The content is scientifically correct.

Agree

10. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chosen, otherwise write "n/a")

n/a

11. Novelty of the paper: The paper provides novel methods, applications, findings or results. Please do not narrowly view "novelty" as only new methods or theories. Papers proposing novel musical applications of existing methods from other research fields are considered novel at ISMIR conferences.

Agree

12. The paper provides all the necessary details or material to reproduce the results described in the paper. Keep in mind that ISMIR respects the diversity of academic disciplines, backgrounds, and approaches. Although ISMIR has a tradition of publishing open datasets and open-source projects to enhance the scientific reproducibility, ISMIR accepts submissions using proprietary datasets and implementations that are not sharable. Please do not simply reject the paper when proprietary datasets or implementations are used.

Agree

13. Pioneering proposals: This paper proposes a novel topic, task or application. Since this is intended to encourage brave new ideas and challenges, papers rated "Strongly Agree" and "Agree" can be highlighted, but please do not penalize papers rated "Disagree" or "Strongly Disagree". Keep in mind that it is often difficult to provide baseline comparisons for novel topics, tasks, or applications. If you think that the novelty is high but the evaluation is weak, please do not simply reject the paper but carefully assess the value of the paper for the community.

Disagree (Standard topic, task, or application)

14. Reusable insights: The paper provides reusable insights (i.e. the capacity to gain an accurate and deep understanding). Such insights may go beyond the scope of the paper, domain or application, in order to build up consistent knowledge across the MIR community.

Agree

15. Please explain your assessment of reusable insights in the paper.

The annotated dataset could be a valuable resource for MIR research, especially for controllable symbolic generation. However, the quality of the generated music remains limited, which may reduce its utility for training or evaluating other MIR models relying on musical coherence or expressivity.

16. Write ONE line (in your own words) with the main take-home message from the paper.

A preliminary study based on a LLM for enhancing music generation from text prompt.

19. Potential to generate discourse: The paper will generate discourse at the ISMIR conference or have a large influence/impact on the future of the ISMIR community.
Agree

20. Overall evaluation: Keep in mind that minor flaws can be corrected, and should not be a reason to reject a paper. Please familiarize yourself with the reviewer guidelines at https://ismir.net/reviewer-guidelines
Weak accept

21. Main review and comments for the authors. Please summarize strengths and weaknesses of the paper. It is essential that you justify the reason for the overall evaluation score in detail. Keep in mind that belittling or sarcastic comments are not appropriate.
This paper investigates the use of a large language model (LLM), specifically the Bloom model [19], for generating captions that are then used to create music from text via the MST-Text model.

Strenghts:
-The experiment includes a listening evaluation conducted by expert musicians, which adds credibility to the results.
-The experiments are well-described and show some improvements offered by the proposed approach.

Weaknesses:
-The experiments are well-described and clearly demonstrate the improvements offered by the proposed approach.

-The generated music lacks naturalness and does not sound as convincing as other existing systems (e.g., https://musicgeneratorai.com/).

-The MST-Text method seems to be a minor modification of the All-MiniLM-L6-v2 model [22], but the details provided are insufficient to fully understand the extent of the modification.

Recommendation:
Despite these weaknesses, I recommend accepting this paper as it introduces new ideas that could pave the way for future research on the application of large language models (LLMs) to Music Information Retrieval (MIR).

Reviewer #2

Questions

2. I am an expert on the topic of the paper.

Agree

3. The title and abstract reflect the content of the paper.

Agree

4. The paper discusses, cites and compares with all relevant related work

Agree

5. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chosen, otherwise write "n/a")

n/a

6. Readability and paper organization: The writing and language are clear and structured in a logical manner.

Agree

7. The paper adheres to ISMIR 2025 submission guidelines (uses the ISMIR 2025 template, has at most 6 pages of technical content followed by "n" pages of references or ethical considerations, references are well formatted). If you selected "No", please explain the issue in your comments.

Yes

8. Relevance of the topic to ISMIR: The topic of the paper is relevant to the ISMIR community. Note that submissions of novel music-related topics, tasks, and applications are highly encouraged. If you think that the paper has merit but does not exactly match the topics of ISMIR, please do not simply reject the paper but instead communicate this to the Program Committee Chairs. Please do not penalize the paper when the proposed method can also be applied to non-music domains if it is shown to be useful in music domains.

Agree

9. Scholarly/scientific quality: The content is scientifically correct.

Agree

10. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chosen, otherwise write "n/a")

n/a

11. Novelty of the paper: The paper provides novel methods, applications, findings or results. Please do not narrowly view "novelty" as only new methods or theories. Papers proposing novel musical applications of existing methods from other research fields are considered novel at ISMIR conferences.

Agree

12. The paper provides all the necessary details or material to reproduce the results described in the paper. Keep in mind that ISMIR respects the diversity of academic disciplines, backgrounds, and approaches. Although ISMIR has a tradition of publishing open datasets and open-source projects to enhance the scientific reproducibility, ISMIR accepts submissions using proprietary datasets and implementations that are not sharable. Please do not simply reject the paper when proprietary datasets or implementations are used.

Agree

13. Pioneering proposals: This paper proposes a novel topic, task or application. Since this is intended to encourage brave new ideas and challenges, papers rated "Strongly Agree" and "Agree" can be highlighted, but please do not penalize papers rated "Disagree" or "Strongly Disagree". Keep in mind that it is often difficult to provide baseline comparisons for novel topics, tasks, or applications. If you think that the novelty is high but the evaluation is weak, please do not simply reject the paper but carefully assess the value of the paper for the community.

Disagree (Standard topic, task, or application)

14. Reusable insights: The paper provides reusable insights (i.e. the capacity to gain an accurate and deep understanding). Such insights may go beyond the scope of the paper, domain or application, in order to build up consistent knowledge across the MIR community.

Agree

15. Please explain your assessment of reusable insights in the paper.

The authors present MetaScore, a dataset with over 900k musical scores paired with textual metadata, collected from the MuseScore forum. The dataset could be used to foster new works regarding text-to-symbolic music generation.

16. Write ONE line (in your own words) with the main take-home message from the paper.

The paper presents MetaScore, a new dataset of symbolic music paired with textual descriptors, and a novel model for text-to-symbolic music generation.

19. Potential to generate discourse: The paper will generate discourse at the ISMIR conference or have a large influence/impact on the future of the ISMIR community.

Agree

20. Overall evaluation: Keep in mind that minor flaws can be corrected, and should not be a reason to reject a paper. Please familiarize yourself with the reviewer guidelines at https://ismir.net/reviewer-guidelines

Weak accept

21. Main review and comments for the authors. Please summarize strengths and weaknesses of the paper. It is essential that you justify the reason for the overall

evaluation score in detail. Keep in mind that belittling or sarcastic comments are not appropriate.

The authors present MetaScore, a dataset with over 900k musical scores paired with textual metadata, collected from the MuseScore forum. A text-conditioned symbolic music generation model is proposed, controlling instruments, genre, composer, complexity and other features.

The background coverage is appropriate, although the inclusion of other relevant symbolic music datasets could have strengthen it (e.g. DadaGP dataset, GigaMIDI dataset). However, the contribution from MetaScore is clear.

The dataset collection, supported by Figure 2., seems consistent. A few doubts persist, after reading the section and inspecting the plots. For example, `Rock` seems to be, together with `Pop`, well represented in the dataset. However, the `guitar`, which I'd argue is heavily used in `Rock`, seems to be missing from the most common 10 instruments. Furthermore, the authors state that "We also standardize the names of well-known musicians to their full
names; for instance, "mozart" is changed to "wolfgang amadeus mozart." - it would be beneficial if a few sentences on why this was done could be presented. Lines 173 to 184 are redundant, as the authors explain the idea behind the MMT genre tagger twice - please correct this. Regarding the choice of the 8 genre tags, although the authors seem to be aware of this as a limitation, the inclusion of Pop with `Rock & Metal` seems to me inadequate, specially when in the `MetaScore-Genre` (presented in Table2) there seems to be a similar number of `Rock` and `Pop` songs.

In 5. the methodology seems appropriate. However, it is a bit unclear how the authors tackle the inclusion of genre, instrument, complexity and composer information. A good reference to include could be GTR-CTRL, in which the authors condition symbolic music generation with genre and instrumentation-specific tokens.

Regarding 6, it would be interesting to see if the results presented in 6.3 have statistical significance - this could strengthen the work.

Finally, the ethical approach of releasing only the files and metadata in the public domain is commendable.

The supporting github page is very well structured and positively impacts the paper.

Minor remarks and nitpicks:

- "To approach text-to-music generation, we leverage a pre-trained large language model (LLM) to generate pseudo
natural language captions from the metadata." - it's not clear in the abstract if the authors used an LLM to get the metadata captions, or to generate music. Maybe rephrase if possible.
- For Figure 2., despite it's relevance, I'd invite the authors to augment the fonts of the labels, in order to faciliate visualization. The tilted Y label (e.g. Copyright Level, Genre, Time Signature) could easily be dropped in order to save space, as they can be inferred from the plot name easily.
- In 3.1. - "We scrapped..." - I believe the right term is "scraped".
- I'm assuming in 3.3. that `MetaScore-Genre` is the subset of `MetaScore-Raw` that has native genre annotations - if that is the case, please clarify. It is specified in 4., but not at this point.
- Footnote 5 - is "copyright" suppose to be there?
- In 8 - only "used" for research purposes (instead of "use").

Reviewer #3

Questions

2. I am an expert on the topic of the paper.
Strongly agree

3. The title and abstract reflect the content of the paper.
Strongly agree

4. The paper discusses, cites and compares with all relevant related work
Disagree

5. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chosen, otherwise write "n/a")
n/a

6. Readability and paper organization: The writing and language are clear and structured in a logical manner.
Strongly agree

7. The paper adheres to ISMIR 2025 submission guidelines (uses the ISMIR 2025 template, has at most 6 pages of technical content followed by "n" pages of references or ethical considerations, references are well formatted). If you selected "No", please explain the issue in your comments.
Yes

8. Relevance of the topic to ISMIR: The topic of the paper is relevant to the ISMIR community. Note that submissions of novel music-related topics, tasks, and

applications are highly encouraged. If you think that the paper has merit but does not exactly match the topics of ISMIR, please do not simply reject the paper but instead communicate this to the Program Committee Chairs. Please do not penalize the paper when the proposed method can also be applied to non-music domains if it is shown to be useful in music domains.

Strongly agree

9. Scholarly/scientific quality: The content is scientifically correct.

Agree

10. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chosen, otherwise write "n/a")

n/a

11. Novelty of the paper: The paper provides novel methods, applications, findings or results. Please do not narrowly view "novelty" as only new methods or theories. Papers proposing novel musical applications of existing methods from other research fields are considered novel at ISMIR conferences.

Agree

12. The paper provides all the necessary details or material to reproduce the results described in the paper. Keep in mind that ISMIR respects the diversity of academic disciplines, backgrounds, and approaches. Although ISMIR has a tradition of publishing open datasets and open-source projects to enhance the scientific reproducibility, ISMIR accepts submissions using proprietary datasets and implementations that are not sharable. Please do not simply reject the paper when proprietary datasets or implementations are used.

Agree

13. Pioneering proposals: This paper proposes a novel topic, task or application. Since this is intended to encourage brave new ideas and challenges, papers rated "Strongly Agree" and "Agree" can be highlighted, but please do not penalize papers rated "Disagree" or "Strongly Disagree". Keep in mind that it is often difficult to provide baseline comparisons for novel topics, tasks, or applications. If you think that the novelty is high but the evaluation is weak, please do not simply reject the paper but carefully assess the value of the paper for the community.

Disagree (Standard topic, task, or application)

14. Reusable insights: The paper provides reusable insights (i.e. the capacity to gain an accurate and deep understanding). Such insights may go beyond the scope of the paper, domain or application, in order to build up consistent knowledge across the MIR community.

Agree

15. Please explain your assessment of reusable insights in the paper.

MetaScore offers reusable insights for MIR and multimodal AI research, fostering consistent knowledge development across these fields.

16. Write ONE line (in your own words) with the main take-home message from the paper.

MetaScore introduces a large-scale symbolic music dataset with captions, enabling controllable music generation.

19. Potential to generate discourse: The paper will generate discourse at the ISMIR conference or have a large influence/impact on the future of the ISMIR community.

Disagree

20. Overall evaluation: Keep in mind that minor flaws can be corrected, and should not be a reason to reject a paper. Please familiarize yourself with the reviewer guidelines at https://ismir.net/reviewer-guidelines

Weak accept

21. Main review and comments for the authors. Please summarize strengths and weaknesses of the paper. It is essential that you justify the reason for the overall evaluation score in detail. Keep in mind that belittling or sarcastic comments are not appropriate.

This paper significantly advances symbolic music generation through MetaScore, a large-scale, publicly accessible musicXML dataset, and innovative text- and tag-based models. The dataset's extensive metadata and the models' controllability offer reusable insights for MIR and multimodal AI research, enhancing user-centric music creation. Despite minor technical ambiguities, the work is impactful and acceptable for its contributions and broad applicability. I strongly hope that the dataset is made publicly available prior to the paper's publication.

I lowered the score because I rediscovered a significant flaw in this paper:
-"Similar to MMT [15], we decompose note-on events to beat and position to reduce the size of the vocabulary and to help the model learn the rhythmic structure of music."--- If this is the main goal of the paper, then the authors should have outperformed text2midi in terms of groove consistency which isn't the case. Also, this authors didn't report on any structure related metrics like Compression ratio as in the text2midi paper. It brings into question how this model performs structurally.

Raised questions:
- Key Signature Extraction: The paper states that key signatures were extracted from MuseScore files, yet MuseScore is known not to explicitly distinguish between relative keys (e.g., G major and E minor). How were G major and E minor separately identified in Fig. 2's data statistics? The authors are requested to provide an explanation (e.g. any preprocessing or analytical techniques applied).
- Key Distribution Imbalance: Significant underrepresentation of common keys like C

major, as observed in Fig. 2. How does this imbalance impact the generalizability of the proposed models, and were any augmentation strategies considered?
- Citation [1]: I wasn't able to find the paper {F. Bhandari and C. Others, "Text2midi: Generating symbolic music," Journal of Music and AI, vol. 1, pp. 100–110, 2024}. DId you mean {Bhandari, Keshav, et al. "Text2midi: Generating Symbolic Music from Captions." Proceedings of the AAAI Conference on Artificial Intelligence, vol. 39, no. 22, 2025}?

View Meta-Reviews

Paper ID
32

Paper Title
Generating Symbolic Music from Natural Language Prompts using an LLM-Enhanced Dataset

Track Name
Papers

2. I am an expert on the topic of the paper.
Agree

3. The title and abstract reflect the content of the paper.
Agree

4. The paper discusses, cites and compares with all relevant related work.
Agree

5. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chosen, otherwise write "n/a")
n/a

6. Readability and paper organization: The writing and language are clear and structured in a logical manner.
Agree

7. The paper adheres to ISMIR 2025 submission guidelines (uses the ISMIR 2025 template, has at most 6 pages of technical content followed by "n" pages of references or ethical considerations, references are well formatted). If you selected "No", please explain the issue in your comments.
Yes

8. Relevance of the topic to ISMIR: The topic of the paper is relevant to the ISMIR community. Note that submissions of novel music-related topics, tasks, and applications are highly encouraged. If you think that the paper has merit but does not exactly match the topics of ISMIR, please do not simply reject the paper but instead communicate this to the Program Committee Chairs. Please do not penalize the paper when the proposed method can also be applied to non-music domains if it is shown to be useful in music domains.

Strongly agree

9. Scholarly/scientific quality: The content is scientifically correct.

Agree

10. Please justify the previous choice (Required if "Strongly Disagree" or "Disagree" is chose, otherwise write "n/a")

n/a

11. Novelty of the paper: The paper provides novel methods, applications, findings or results. Please do not narrowly view "novelty" as only new methods or theories. Papers proposing novel musical applications of existing methods from other research fields are considered novel at ISMIR conferences.

Disagree

12. The paper provides all the necessary details or material to reproduce the results described in the paper. Keep in mind that ISMIR respects the diversity of academic disciplines, backgrounds, and approaches. Although ISMIR has a tradition of publishing open datasets and open-source projects to enhance the scientific reproducibility, ISMIR accepts submissions using proprietary datasets and implementations that are not sharable. Please do not simply reject the paper when proprietary datasets or implementations are used.

Agree

13. Pioneering proposals: This paper proposes a novel topic, task or application. Since this is intended to encourage brave new ideas and challenges, papers rated "Strongly Agree" and "Agree" can be highlighted, but please do not penalize papers rated "Disagree" or "Strongly Disagree". Keep in mind that it is often difficult to provide baseline comparisons for novel topics, tasks, or applications. If you think that the novelty is high but the evaluation is weak, please do not simply reject the paper but carefully assess the value of the paper for the community.

Disagree (Standard topic, task, or application)

14. Reusable insights: The paper provides reusable insights (i.e. the capacity to gain an accurate and deep understanding). Such insights may go beyond the scope of the paper, domain or application, in order to build up consistent knowledge across the MIR community.

Agree

15. Please explain your assessment of reusable insights in the paper.

The paper offers insights into how LLMs can be effectively leveraged to enhance music datasets and improve text-to-music generation. Generating pseudo-captions from metadata and the comparative analysis of tag-based vs. text-based control provide some insights for future research in this area.

The dataset will also be useful for several tasks.

16. Write ONE line (in your own words) with the main take-home message from the paper.

This paper introduces MetaScore, a large-scale symbolic music dataset enhanced by LLM-generated captions, enabling improved text-to-music generation with versatile controls.

19. Potential to generate discourse: The paper will generate discourse at the ISMIR conference or have a large influence/impact on the future of the ISMIR community.

Agree

20. Overall evaluation (to be completed before the discussion phase): Please first evaluate before the discussion phase. Keep in mind that minor flaws can be corrected, and should not be a reason to reject a paper. Please familiarize yourself with the reviewer guidelines at https://ismir.net/reviewer-guidelines.

Weak accept

21. Main review and comments for the authors (to be completed before the discussion phase). Please summarize strengths and weaknesses of the paper. It is essential that you justify the reason for the overall evaluation score in detail. Keep in mind that belittling or sarcastic comments are not appropriate.

This paper introduces MetaScore, a new large-scale dataset of 963K musical scores with rich metadata. An LLM is used to generate pseudo-captions from the metadata, and text-to-music and tag-to-music models are trained on this dataset.

The paper is well-written and clearly structured, and has several strengths:

It introduces a valuable dataset, addressing a significant gap in the availability of large-scale, richly annotated symbolic music datasets.The data collection method (scraping) raises ethical considerations, which the authors attempt to address by only releasing public domain ones (while those not in public domain will be shared upon request and only for research purposes). Beyond the data collection, the authors also contribute through several dataset enhancements, including: the extraction and standardization of metadata, including key signature, time signature, tempo, and instrument information, with a focus on General MIDI compatibility and consistency in composer names. Then, to address missing genre information, a genre tagger is trained, enabling genre-controlled music generation. The accuracy of this tagger is evaluated through objective and subjective tests. Finally, LLMs are leveraged to generate pseudo-captions from the metadata, creating text descriptions of the music that facilitate text-to-music generation. Then a text-conditioned music generation model is proposed, which can be controlled using different musical properties (e.g. genre, instruments, composer, etc...)

The paper also presents a comprehensive evaluation, including both objective and subjective measures, and compares the proposed models with relevant baselines, and the results demonstrate the effectiveness of the proposed approach for controllable symbolic music generation.

In order to improve the paper there are a few things which could be clarified.

It is mentioned that: "We also standardize the names of well-known musicians to their full names". How is this actually done? What is the reference for the composer's full name? It would

be helpful to know the specific resources or authority used for this standardization (e.g., a music database, etc.).

What % of the dataset was excluded due to genres with scarce presence? This would give a better understanding of how much data was deemed unusable and the potential impact on the diversity of the generated music. Would it be possible to assign these genres to broader categories?

What are the details of the split size in the genre tagger? E.g. size of the validation split

In the objective evaluation: values closer to the ground truth indicate better performance. While these objective metrics are valuable, it would be beneficial to also assess the originality of the generated music to ensure the model is not simply reproducing segments from the training data. Metrics or techniques to measure novelty or dissimilarity from the training set could be considered.

In terms of formatting and grammatical errors, please consider the following:

Grammar/Typos: "scrapped" → "scraped" (line 133); "purpose → purpose" (line 456)

Clarity: "we decompose note-on events to beat and position" (line 279) - clarify "beat number and position within beat";

"The description of "six special structural events" (lines 308-314) is wordy."

Repetition: "Lines 182-184 are repeated; 332-335 was already mentioned before"

22. Final recommendation (to be completed after the discussion phase) Please give a final recommendation after the discussion phase. In the final recommendation, please do not simply average the scores of the reviewers. Note that the number of recommendation options for reviewers is different from the number of options here. We encourage you to take a stand, and preferably avoid "weak accepts" or "weak rejects" if possible.
Accept

23. Meta-review and final comments for authors (to be completed after the discussion phase)
Dear Authors,

Thank you for your contribution to ISMIR 2025 with your paper titled "Generating Symbolic Music from Natural Language Prompts using an LLM-Enhanced Dataset. Below is a summary of the reviews provided by the reviewers and the meta-reviewer, with some suggestions for improvement.

This paper introduces MetaScore, a new large-scale dataset of 963K musical scores with rich metadata. An LLM is used to generate pseudo-captions from the metadata, and text-to-music and tag-to-music models are trained on this dataset.

Summary of strengths

All reviewers (R1, R2, R3, Meta) noted MetaScore as a significant dataset contribution, filling a

crucial gap in richly annotated symbolic music datasets. The ethical considerations in data collection, by releasing only public domain files and sharing others for research upon request, were also commended (Meta, R2). The paper's use of LLMs to generate pseudo-captions from metadata was highlighted as a key strength, providing reusable insights for future research in leveraging LLMs for music datasets. The proposed text-conditioned music generation model, enabling controllable music generation through various musical properties, was also seen as a notable advancement (Meta, R2). Reviewers found the paper well-written and clearly structured (R1, R2, R3, Meta) and its topic highly relevant to the ISMIR community (R1, R2, R3, Meta). The inclusion of comprehensive objective and subjective evaluations, including expert musician listening tests, added credibility (Meta, R1).

Summary of weaknesses

Novelty concerns: mixed opinions on the paper's novelty, with the task itself being considered standard
Insufficient technical detail: several aspects of the methodology, including data standardization, genre classification, and specific model modifications, lacked sufficient detail.
Presentation and clarity issues: the paper contained minor formatting errors, typos, redundant phrasing, and occasional lack of clarity in descriptions.
Evaluation nuances: while comprehensive, there were suggestions for incorporating discussions on the originality of generated music and statistical significance.

Here some suggestions to address the weaknesses for the camera-ready version, but please also consider each of the individual reviews:

Clarify technical details:
Briefly explain the method or reference used for standardizing composer names (e.g., specific music database, manual curation) (Meta, R2).
State the percentage of the dataset excluded due to genres with scarce presence (Meta, R2).
Specify the validation split size for the genre tagger (Meta).
Provide clearer explanations of several aspects mentioned in each of the reviews, e.g. "beat number and position within beat", the modifications made to the All-MiniLM-L6-v2 model for MST-Text (R1).
Briefly acknowledge how the distinction between relative keys (e.g., G major and E minor) was handled during key signature extraction (R3).

Improve presentation and clarity issues:
Proofread meticulously for any remaining typos and grammatical errors (e.g., "scrapped" to "scraped)

Enhance the explanation of the "six special structural events" (Meta).

Remove redundant lines of text (e.g., lines 182-184 and 332-335) (Meta).

Consider enhancing Figure 2 readability by augmenting font sizes for labels and considering the removal of tilted Y-labels where redundant (R2).

Verify the citation for "Text2midi" (R3).

Address evaluation concerns

In the limitations section, briefly discuss the originality of the generated music and acknowledge this as a future research direction (Meta, R1).

A brief mention of statistical significance for the results presented in Section 6.3 could strengthen the work (R2)

Overall, we believe that this article will constitute a significant contribution to ISMIR 2025, and these revisions and improvements will enhance the paper's clarity, depth, and impact.

Best regards,
Meta Reviewer