Ethics Statements in AI Music Papers: The Effective and the Ineffective

Julia Barnett
Northwestern University

Patrick O'Reilly
Northwestern University

Jason Brent Smith
Northwestern University

Annie Chu
Northwestern University

Bryan Pardo
Northwestern University

Abstract

While research in AI methods for music generation and analysis has grown in scope and impact, AI researchers' engagement with the ethical consequences of this work has not kept pace. To encourage such engagement, many publication venues have introduced optional or required ethics statements for AI research papers. Though some authors use these ethics statements to critically engage with the broader implications of their research, we find that the majority of ethics statements in the AI music literature do not appear to be effectively utilized for this purpose. In this work, we conduct a review of ethics statements across ISMIR, NIME, and selected prominent works in AI music from the past five years. We then offer suggestions for both audio conferences and researchers for engaging with ethics statements in ways that foster meaningful reflection rather than formulaic compliance.

1 Introduction

AI music and audio models are gaining more traction and prominence in society every day. Conferences like the International Society for Music Information Retrieval (ISMIR) and the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) have cited record high submissions over the past couple years. Text-to-music commercial generative models like Suno and Udio report tens of millions of users [61]. Big Tech firms like Google [2, 59] and Meta [15] continue to release their own music models. What does *not* continue to grow at the same rate is an internal reflection on ethical issues of these models from the creators of the works.

In recent years, computer science researchers have sought to address, engage with, and mitigate the negative impacts of their work through a change to the peer review process [25]. NeurIPS even implemented a mandatory broader impact statement in 2020 [36], but it is now defunct and instead baked into a smaller aspect of a required checklist. Conferences such as ISMIR have recently begun implementing an optional ethics statement that does not count towards the word limit, citing "the goal of encouraging more discussion of ethical considerations in the ISMIR research community."[1] This is in line with a larger trend of conferences encouraging ethical reflection by allowing or requiring an extra page for ethics and broader impact statements, e.g., the International Conference on Machine Learning (ICML) and the ACM Conference on Fairness, Accountability, and Transparency (FAccT).[2]

These statements are a valuable opportunity for authors and creators of these models to engage with the ethical implications of their work. However, when optional, only a minority of people choose to use them: for instance, only 28% of papers published at ISMIR in 2024 used the page, and often they are not used for the designated purpose. This is reflective of a wider issue: a recent review of

[1] https://ismir2024.ismir.net/authors/call-for-papers

[2] https://icml.cc/Conferences/2025/CallForPapers, https://facctconference.org/2025/aguide

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: AI for Music.

generative audio works found that less than 10% of papers discuss any potential negative impacts of their work [4]. This is not indicative of a lack of potential harm—the work also details a suite of harms including climate impact, copyright infringement, deepfakes, and creativity stifling—but rather a failure by the research community to ethically engage with their own work.

We see this moment as an opportunity: as ethics statements are increasingly incorporated into the research process, there is space to support researchers in approaching them productively. We seek to contribute to that effort by offering guidance, examples, and suggestions for engaging with ethics statements in ways that foster meaningful reflection rather than formulaic compliance. In this work, we analyze how authors in music-AI research are currently using these ethics statements, outline key categories of potential harm that may warrant discussion, and provide calls to action to both conferences and researchers. The goal of this work is not to encourage formulaic statements, but to support authors in developing thoughtful, substantive accounts of ethical implications of their work in order to become better scientists.

2 Current State of Ethics in AI Music: A Mix of the Good, Bad, and Ugly

2.1 ISMIR 2024: An Optional Bonus Page

The optional extra page ISMIR provides which we describe above (Sec. 1) gives researchers an option to engage with the broader impact and ethical implications of their work without having to cut any content from the main body of the paper. Even with this added bonus page option, only $n = 37/133(28\%)$ of works wrote one. Further, the average statement was 1.8 paragraphs long, and only 169 words long (median of 148)—relative to the larger paper, these are very short. Of these 37 ethics statements, 6 (16%) limited their ethical discussion to a statement that their project was approved by IRB. Two $(n = 2; 5\%)$ other works used this space not to engage ethically with their work, but to state they did not see any potential for ethical issues or to describe the history and ownership of the data they were using, without further elaboration of the ethical impact of their own work. That leaves $29/133(22\%)$ papers published in ISMIR proceedings that detailed potential ethical implications of their work using this bonus page.

Across these 29 ethics statements, there were 17 unique harms discussed (detailed in the Appendix in Table 1). The most prominent was the potential for copyright infringement $(n = 13; 45\%)$ through use of released datasets or models trained on those datasets. The second most discussed was the potential for labor displacement $(n = 8; 28\%)$ as a result of these tools being used in industry in place of artists or other music and sound-related professions such as sound engineers. Tied for third $(n = 7; 24\%)$ were general unspecified bias of the training data and the potential for misappropriation of cultures through unknowingly using these tools. 5 (17%) papers each discussed the potential harm for voice replication or impersonation through misuse of these models, as well as a Western bias of music, sounds, and voices in the training datasets. Only three papers (10%) mentioned the climate impact, which is likely relevant for the majority of these works. Similarly, only three works each acknowledged data scraping and privacy as ethical issues. All other harms were mentioned only once or twice and include sustainability of business models, questions of authorship in AI creations, stifling of creativity, homogenization of music, memorization of training data, loss of agency, replicability of scientific works, and the potential for false positives in a deepfake detection algorithm.

2.2 NIME 2024/2025: A Mandatory Statement in the Main Text

The International Conference on New Interfaces for Musical Expression (NIME) collects research related to musical interface design as well as artistic performances using new musical interfaces. The conference maintains a set of principles and a code of practice for maintaining ethical research [35]. It also includes a strict requirement for ethics statements [3], using the following text to describe guidelines for what authors should include:

> "...information regarding sources of funding, potential conflicts of interest (financial or non-financial), informed consent if the research involved human participants, statement on welfare of animals if the research involved animals or any other information or context that helps ethically situate your research."

[3] https://github.com/NIME-conference/nime-template/releases/tag/v2024.12.02

Of the 50 papers in the 2024 and 2025 NIME proceedings containing the phrases "artificial intelligence," "machine learning," or "neural networks," 49 papers included the required ethics statements. However, without a dedicated page, many of the ethics statements stand as a single paragraph conforming to the exact wording of the guidelines, with few delving deeper than funding and recruitment sources. Themes among these ethics statements include the use of the author's own music as data and the use of a system as a tool specifically designed for the author. More generally, these ethics sections include conflicts of interest, participant informed consent, and ethical data sourcing (e.g., creating one's own datasets or using appropriately licensed material). More detailed reflections tend to arise in projects involving tangible hardware, where sustainability and materials are emphasized, or in practice-based collaborations with artists, where authorship and participation are foregrounded.

Exceptions to this pattern occur in NIME papers related to the philosophical aspects of sustainability and practice in musical interfaces. For example, an analysis of the challenges of preserving deep generative models for NIMEs uses its Ethical Standards section to articulate the paper's goals of minimizing energy usage when recreating models, as well as to consider the authors' inclusion of the original authors of the systems compared [13].

It is important to note that NIME proceedings are diverse, spanning augmented instruments, installations, neural audio plugins, HCI analyses, and human–AI co-creation systems to name a few. Compared to model- and system-driven venues like ICASSP or ISMIR, NIME papers often center specific users (e.g., musicians, performers), practice-based methods, and artistic reflections. As a result, ethical considerations are frequently discussed outside of the ethics section, either in separately dedicated sections of the paper ([3, 31, 52]) or in the discussion/conclusion sections ([47]).

2.3 Prominent Music AI Works of the 2020s: A Mix of Everything

In addition to these two music conferences, we also surveyed prominent music generation papers from the last five years starting with Jukebox [17] (2020) as the watershed moment for AI music generation research. These papers were mainly produced by established academic research groups, Big Tech, or collaborations between the two. We used as a starting point the 16 music generative modeling papers recently assessed for openness by MusGO [6], which cover "a diverse set of music-generative models" and "a variety of architectures commonly used within the field." We bolstered this set with 9 additional papers identified by our team as high-impact or otherwise representative of state-of-the-art methods, none of which overlapped with 2024 ISMIR or NIME. The final set can be found in Table 2. Of these 25 works, $n = 13; (52\%)$ had some form of ethics statement. These were typically a short paragraph or section after the conclusion $(n = 10; 77\%)$. Two were written as part of the discussion, and one was written in the appendix. On average these were slightly longer than ISMIR statements (1.9 paragraphs and 281 words), though this was driven by one much longer and more effective ethics statement [60] (11 paragraphs and 1,576 words); the median was on par with ISMIR at 147 words.

Many of these ethics statements reflect a trend in which researchers seek to preemptively defend their work from criticism rather than to acknowledge and reckon with potential harms. None of these 25 papers mentioned the potential impact on the environment. Only two works ($n = 2$) in this corpus discussed the potential for copyright infringement [51, 60], which is a stark contrast to the 45% of ISMIR ethics statements that discussed this. It is potentially indicative of industry wariness to even acknowledge the potential for copyright infringement, seeing as $n = 20/25(75\%)$ of these works had at least one industry author and $n = 16/25(64\%)$ had a majority of industry affiliated authors. Thickstun et al. [60] assert "we are certain that generative models of music are similarly capable of plagiarism," citing works on memorization in text [11] and vision [55], which has now been echoed by works in music [5, 7]. Works trained on massive amounts of data such as VampNet [24] and Moûsai [51] notably lack this acknowledgment in their papers. VampNet used hundreds of thousands of scraped songs for training data and had no ethics statement, and Moûsai scraped roughly 50,000 songs from Spotify—though the authors disclose their data acquisition process and open-source their scraping code. While this transparency is laudable, the authors do not use the ethics statement to engage with the implications of their use of copyrighted data, but rather to defend this use as technically legal or "exempted from copyright infringement."

About half of these statements focused on stating in explicit terms that they acquired their data through legal means ($n = 6; 46\%$), sometimes without engaging critically beyond stating that ($n = 2$). Similarly, two works ($n = 2$) only briefly mentioned that their training data has biases without engaging beyond a vague description of what that means. The most commonly discussed harm

was labor impact on artists, with $n = 7; 54\%$ of statements discussing this [15, 18, 44, 51, 58, 60, 66]. As [58] notes, "generative models could potentially create an unbalanced competitive environment for artists, a problem that is yet to be solved;" this is a harm that could be mentioned in each and every one of the 25 papers surveyed in this section.

Four of these works [15, 27, 58, 60] discussed the potential harm resulting from training on a predominance of Western music data, which Thickstun et al. note "inevitably biases the model's predictions towards infilling completions that follow Western rules of composition" [60]. This is an example of more critically engaging with the biases present in your dataset and discussing potential implications of that, rather than just vaguely gesturing towards the idea of bias.

3 Call to Action: Changes to this Process

This review of the recent use of ethics statements has made two key aspects clear: (1) **ethics statements are underutilized** both in terms of writing them at all and discussing a comprehensive suite of relevant harms, and (2) researchers are frequently **using these statements ineffectively** by not engaging with the ethical implications of their work. We believe there are a series of changes that could right both of these wrongs, but they require collective action by the research community across various stages of the research process.

Audio conferences should start requiring an extra page allotted for ethics statements instead of including them as an optional page (or even not allowing space for them at all). However, these statements should not be a regurgitation of the ethical guidelines provided by the conference. To help avoid this, the call for papers pages should provide excellent examples of ethics statements from previous years (such as Thickstun et al. [60], Batlle-Roca et al. [7], or Barnett et al. [5]), so as to call attention to ethically conducted work and authors who properly engage critically with their work. The ethics statements could also be required as part of the abstract registration a week before the conference in order to encourage authors to write these statements thoughtfully and not just as an afterthought. There could also be a separate set of reviewers for these ethics statements comprised of authors who submit ethics-focused papers who already have domain expertise and ability to assess the quality of ethical reflections; making this evaluation relevant to acceptance decisions would provide a strong incentive for authors to engage more thoughtfully and substantively with ethical concerns.

Researchers in this field should start viewing these statements as an opportunity to begin meaningfully engaging with the broader implications of their own work. Though they typically come at the end of the work in reflection, one should begin thinking about the ethical implications of the work at the start of the research process. In addition to detailing potential harms in ethics statements, one can describe steps taken to actively mitigate potential harms or even harms the work may be actively addressing. If it is unclear where to start, they should review a set of harms covered by other works in the field (such as the ones described in Appendix Tables 1-2 or in a review of harms of these models [4]). They should identify the harms that could be applicable and detail ways in which the work is vulnerable to those harms, may perpetuate those harms, or has even taken action to fight against these harms. Authors should not merely list these off as a checklist, but take the opportunity to engage with them. For instance, if the researcher performed any training runs on GPUs they can estimate the environmental cost of a training run: number of GPUs $\times$ energy output of GPUs (W) $\times$ hours of training time to get an energy cost in kilowatt-hours (kWh); then the author can disclose this for transparency and to encourage their peers to do so as well [26].

4 Conclusion and Future Work

In this work we examined the usage of ethics statements in music AI works across ISMIR, NIME, and other prominent works in the domain. We found that though some authors are using these effectively to critically engage with the broader implications of their work, there is a dearth of usage and even more so proper usage of these statements. We identify harms that have been discussed which are widely applicable, and we present suggestions to improve this process at large. We recognize we can be a positive force in this domain as researchers intimately familiar with both the ethical implications of these works and the domain itself. In the future we plan to release a more interactive, open, and living document guide to assist authors with engaging with these harms. We envision doing so through a decision-tree type mechanism (e.g., did you use training data in your work? If so these are some potential harms.) If you would like to contribute, please reach out to the authors of this work.

Acknowledgments and Disclosure of Funding

This work was supported in part by USA National Science Foundation award numbers 2222369 and 2300633.

References

[1] Joanne Affolter and Martin A Rohrmeier. Utilizing listener-provided tags for music emotion recognition: A data-driven approach. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[2] Andrea Agostinelli, Timo I Denk, Zalán Borsos, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, Aren Jansen, Adam Roberts, Marco Tagliasacchi, et al. Musiclm: Generating music from text. *arXiv preprint arXiv:2301.11325*, 2023.

[3] Jack Armitage, Victor Shepardson, and Thor Magnusson. Tölvera: Composing with basal agencies. pages 282–291, September 2024.

[4] Julia Barnett. The ethical implications of generative audio models: A systematic literature review. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pages 146–161, 2023.

[5] Julia Barnett, Hugo Flores Garcia, and Bryan Pardo. Exploring musical roots: Applying audio embeddings to empower influence attribution for a generative music model. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[6] Roser Batlle-Roca, Laura Ibáñez-Martínez, Xavier Serra, Emilia Gómez, and Martín Rocamora. Musgo: A community-driven framework for assessing openness in music-generative ai. *arXiv preprint arXiv:2507.03599*, 2025.

[7] Roser Batlle-Roca, Wei-Hsiang Liao, Xavier Serra, Yuki Mitsufuji, and Emilia Gómez. Towards assessing data replication in music generation with music similarity metrics on raw audio. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[8] Tim Beyer and Angela Dai. End-to-end piano performance-midi to score conversion with transformers. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[9] Irmak Bukey, Michael Feffer, and Chris Donahue. Just label the repeats for in-the-wild audio-to-score alignment. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[10] Antoine Caillon and Philippe Esling. Rave: A variational autoencoder for fast and high-quality neural audio synthesis. *arXiv preprint arXiv:2111.05011*, 2021.

[11] Nicholas Carlini, Daphne Ippolito, Matthew Jagielski, Katherine Lee, Florian Tramer, and Chiyuan Zhang. Quantifying memorization across neural language models. In *The Eleventh International Conference on Learning Representations*, 2022.

[12] Ke Chen, Yusong Wu, Haohe Liu, Marianna Nezhurina, Taylor Berg-Kirkpatrick, and Shlomo Dubnov. Musicldm: Enhancing novelty in text-to-music generation using beat-synchronous mixup strategies. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1206–1210. IEEE, 2024.

[13] Isaac Clarke, Francesco Ardan Dal Rí, and Raul Masu. Longevity of deep generative models in nime: Challenges and practices for reactivation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 224–230, 2025.

[14] Marco Comunità, Zhi Zhong, Akira Takahashi, Shiqi Yang, Mengjie Zhao, Koichi Saito, Yukara Ikemiya, Takashi Shibuya, Shusuke Takahashi, and Yuki Mitsufuji. Specmaskgit: Masked generative modeling of audio spectrograms for efficient audio synthesis and beyond. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[15] Jade Copet, Felix Kreuk, Itai Gat, Tal Remez, David Kant, Gabriel Synnaeve, Yossi Adi, and Alexandre Défossez. Simple and controllable music generation. *Advances in Neural Information Processing Systems*, 36:47704–47720, 2023.

[16] Dorian Desblancs, Gabriel Meseguer-Brocal, Romain Hennequin, and Manuel Moussallam. From real to cloned singer identification. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[17] Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, and Ilya Sutskever. Jukebox: A generative model for music. *arXiv preprint arXiv:2005.00341*, 2020.

[18] Chris Donahue, Antoine Caillon, Adam Roberts, Ethan Manilow, Philippe Esling, Andrea Agostinelli, Mauro Verzetti, Ian Simon, Olivier Pietquin, Neil Zeghidour, et al. Singsong: Generating musical accompaniments from singing. *arXiv preprint arXiv:2301.12662*, 2023.

[19] Zach Evans, CJ Carr, Josiah Taylor, Scott H Hawley, and Jordi Pons. Fast timing-conditioned latent audio diffusion. In *Forty-first International Conference on Machine Learning*, 2024.

[20] Zach Evans, Julian D Parker, CJ Carr, Zack Zukowski, Josiah Taylor, and Jordi Pons. Long-form music generation with latent diffusion. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[21] Zach Evans, Julian D Parker, CJ Carr, Zack Zukowski, Josiah Taylor, and Jordi Pons. Stable audio open. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2025.

[22] Arthur Flexer. On the validity of employing chatgpt for distant reading of music similarity. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[23] Chenyu Gao, Federico Reuben, and Tom Collins. Variation transformer: New datasets, models, and comparative evaluation for symbolic music variation generation. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[24] Hugo Flores Garcia, Prem Seetharaman, Rithesh Kumar, and Bryan Pardo. Vampnet: Music generation via masked acoustic token modeling. *arXiv preprint arXiv:2307.04686*, 2023.

[25] Brent Hecht, Lauren Wilcox, Jeffrey P Bigham, Johannes Schöning, Ehsan Hoque, Jason Ernst, Yonatan Bisk, Luigi De Russis, Lana Yarosh, Bushra Anjum, et al. It's time to do something: Mitigating the negative impacts of computing through a change to the peer review process. *arXiv preprint arXiv:2112.09544*, 2021.

[26] Andre Holzapfel, Anna-Kaisa Kaila, and Petra Jääskeläinen. Green mir?: Investigating computational cost of recent music-ai research in ismir. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[27] Qingqing Huang, Daniel S Park, Tao Wang, Timo I Denk, Andy Ly, Nanxin Chen, Zhengdong Zhang, Zhishuai Zhang, Jiahui Yu, Christian Frank, et al. Noise2music: Text-conditioned music generation with diffusion models. *arXiv preprint arXiv:2302.03917*, 2023.

[28] Max WY Lam, Qiao Tian, Tang Li, Zongyu Yin, Siyuan Feng, Ming Tu, Yuliang Ji, Rui Xia, Mingbo Ma, Xuchen Song, et al. Efficient neural music generation. *Advances in Neural Information Processing Systems*, 36:17450–17463, 2023.

[29] Gael Le Lan, Bowen Shi, Zhaoheng Ni, Sidd Srinivasan, Anurag Kumar, Brian Ellis, David Kant, Varun Nagaraja, Ernie Chang, Wei-Ning Hsu, et al. High fidelity text-guided music editing via single-stage flow matching. *arXiv preprint arXiv:2407.03648*, 2024.

[30] Julian Lenz and Anirudh Mani. Cadenza: A generative framework for expressive musical ideas and variations. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[31] Yikai Li and Ge Wang. Chai => interactive ai tools in chuck. pages 553–559, September 2024.

[32] Lucas S Maia, Richa Namballa, Martín Rocamora, Magdalena Fuentes, and Carlos Guedes. Investigating time-line-based music traditions with field recordings: a case study of candomblé bell patterns. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[33] Martin E Malandro. Composer's assistant 2: Interactive multi-track midi infilling with fine-grained user control. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[34] Jan Melechovsky, Zixun Guo, Deepanway Ghosal, Navonil Majumder, Dorien Herremans, and Soujanya Poria. Mustango: Toward controllable text-to-music generation. *arXiv preprint arXiv:2311.08355*, 2023.

[35] Fabio Morreale, Nicolas Gold, Cécile Chevalier, and Raul Masu. Nime principles & code of practice on ethical research. 2023.

[36] Priyanka Nanayakkara, Jessica Hullman, and Nicholas Diakopoulos. Unpacking the expressed consequences of ai research in broader impact statements. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 795–806, 2021.

[37] Shahan Nercessian, Johannes Imort, Ninon Devis, and Frederik Blang. Generating sample-based musical instruments using neural audio codec language models. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[38] Ngan VT Nguyen, Elizabeth AM Acosta, Tommy Dang, and David RW Sears. Exploring internet radio across the globe with the mirage online dashboard. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[39] Ziqian Ning, Huakang Chen, Yuepeng Jiang, Chunbo Hao, Guobin Ma, Shuai Wang, Jixun Yao, and Lei Xie. Diffrhythm: Blazingly fast and embarrassingly simple end-to-end full-length song generation with latent diffusion. *arXiv preprint arXiv:2503.01183*, 2025.

[40] Javier Nistal, Marco Pasini, Cyran Aouameur, Maarten Grachten, and Stefan Lattner. Diff-a-riff: Musical accompaniment co-creation via latent diffusion models. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[41] Javier Nistal, Marco Pasini, and Stefan Lattner. Improving musical accompaniment co-creation via diffusion transformers. *arXiv preprint arXiv:2410.23005*, 2024.

[42] Zachary Novack, Zach Evans, Zack Zukowski, Josiah Taylor, CJ Carr, Julian Parker, Adnan Al-Sinan, Gian Marco Iodice, Julian McAuley, Taylor Berg-Kirkpatrick, et al. Fast text-to-audio generation with adversarial post-training. *arXiv preprint arXiv:2505.08175*, 2025.

[43] Zachary Novack, Julian McAuley, Taylor Berg-Kirkpatrick, and Nicholas Bryan. Ditto-2: Distilled diffusion inference-time t-optimization for music generation. *arXiv preprint arXiv:2405.20289*, 2024.

[44] Zachary Novack, Julian McAuley, Taylor Berg-Kirkpatrick, and Nicholas J Bryan. Ditto: Diffusion inference-time t-optimization for music generation. *arXiv preprint arXiv:2401.12179*, 2024.

[45] Julian D Parker, Janne Spijkervet, Katerina Kosta, Furkan Yesiler, Boris Kuznetsov, Ju-Chiang Wang, Matt Avent, Jitong Chen, and Duc Le. Stemgen: A music generation model that listens. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1116–1120. IEEE, 2024.

[46] Vjosa Preniqi, Iacopo Ghinassi, Julia Ive, Kyriaki Kalimeri, and Charalampos Saitis. Automatic detection of moral values in music lyrics. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[47] Nicola Privato, Victor Shepardson, Giacomo Lepri, and Thor Magnusson. Stacco: Exploring the embodied perception of latent representations in neural synthesis. pages 424–431, September 2024.

[48] Pedro Ramoneda, Vsevolod Eremenko, Alexandre D'Hooge, Emilia Parada-Cabaleiro, and Xavier Serra. Towards explainable and interpretable musical difficulty estimation: a parameter-efficient approach. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[49] Pedro Ramoneda, Martin Rocamora, and Taketo Akama. Music proofreading with refinpaint: where and how to modify compositions given context. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[50] Xavier Riley, Zixun Guo, Drew Edwards, and Simon Dixon. Gaps: A large and diverse classical guitar dataset and benchmark transcription model. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[51] Flavio Schneider, Ojasv Kamal, Zhijing Jin, and Bernhard Schölkopf. Mo\^usai: Text-to-music generation with long-context latent diffusion. *arXiv preprint arXiv:2301.11757*, 2023.

[52] Nicholas Shaheed and Ge Wang. I am sitting in a (latent) room. pages 333–338, September 2024.

[53] Adithi Shankar, Genís Plaja-Roglans, Thomas Nuttall, Martín Rocamora, and Xavier Serra. Saraga audiovisual: a large multimodal open data collection for the analysis of carnatic music. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[54] Nithya Shikarpur, Krishna Maneesha Dendukuri, Yusong Wu, Antoine Caillon, and Cheng-Zhi Anna Huang. Hierarchical generative modeling of melodic vocal contours in hindustani classical music. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[55] Gowthami Somepalli, Vasu Singla, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Diffusion art or digital forgery? investigating data replication in diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6048–6058, 2023.

[56] Robert Sowula and Peter Knees. Mosaikbox: Improving fully automatic dj mixing through rule-based stem modification and precise beat-grid estimation. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[57] Hitoshi Suda, Shunsuke Yoshida, Tomohiko Nakamura, Satoru Fukayama, and Jun Ogata. Fruitsmusic: A real-world corpus of japanese idol-group songs. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[58] Or Tal, Alon Ziv, Itai Gat, Felix Kreuk, and Yossi Adi. Joint audio and symbolic conditioning for temporally controlled text-to-music generation. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[59] Lyria Team. Magenta realtime. 2025.

[60] John Thickstun, David Hall, Chris Donahue, and Percy Liang. Anticipatory music transformer. *arXiv preprint arXiv:2306.08620*, 2023.

[61] UMG Recordings, Inc. v. Suno, Inc. (2025); Main complaint filed 06/24/2024.

[62] Soumya Sai Vanka, Christian Steinmetz, Jean-Baptiste Rolland, Joshua Reiss, and George Fazekas. Diff-mst: Differentiable mixing style transfer. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[63] Karn N Watcharasupat and Alexander Lerch. A stem-agnostic single-decoder system for music source separation beyond four stems. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[64] Benno Weck, Ilaria Manco, Emmanouil Benetos, Elio Quinton, George Fazekas, and Dmitry Bogdanov. Muchomusic: Evaluating music understanding in multimodal audio-language models. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[65] Shangda Wu, Yashan Wang, Xiaobing Li, Feng Yu, and Maosong Sun. Melodyt5: A unified score-to-score transformer for symbolic music processing. In *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)*, 2024.

[66] Shih-Lun Wu, Chris Donahue, Shinji Watanabe, and Nicholas J Bryan. Music controlnet: Multiple time-varying controls for music generation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 32:2692–2703, 2024.

A Ethics Statement

In this section we write our own ethics statement for this work, demonstrating both what we believe to be an effective ethics statement (A.1), which engages with the specific potential harms related to the content of the work, and an ineffective ethics statement (A.2), which briefly addresses "boilerplate" issues unrelated to the potential harms of this particular work.

A.1 An Effective Ethics Statement

The primary concern we have with writing this work is that though we encourage ethical engagement with work, researchers may use our suggestions for harms to discuss in music AI papers as boxes to check in writing a formulaic ethics statement. As a result, though we would be encouraging more researchers to write ethics statements, we would be inadvertently contributing to the exact problem identified in this work which is ineffective usage of these statements.

Another concern we have with this work is that our sample is biased in multiple ways. With regard to ISMIR and NIME, we focused only on the ethics statements and did not include the main body of text in the corpus which we analyzed—as a result, we may be categorizing works that critically engaged with the broader implications of their work in the main text as not doing so in the manner we focused on. The papers we selected to reflect a sample of 25 prominent music AI papers in the last five years was certainly sampled to those we believe to be "prominent," and that is inherently biased to our own views. We attempted to circumvent this by basing the bulk of these works on those selected by [6] because we both agreed with the selection and it was then verified my two independent sets of authors, but this is a biased sample nonetheless.

Finally we acknowledge that calling to attention instances where researchers did not write an ethics statement may make some authors of these works uneasy, and we are not attempting to shame anyone by doing so. We are focusing on works that are already prominent in the field in terms of citations or big names (e.g., Google and Meta), knowing this work will not have a deleterious impact on them so as to not unjustly critique any nascent works. By doing so we highlight that all works in music AI going forward should have ethics statements and these big names are no exception. Nevertheless we acknowledge that this could be viewed in a negative light by our peers, and reiterate causing discomfort was not our intent.

We believe these outlined risks are far outweighed by the potential benefit we view of conducting this review.

A.2 An Ineffective Ethics Statement

We obtained all of the papers analyzed in this work through legal means because they were published on conference proceedings or arXiv. There is likely bias in the sample we analyzed, and future work should seek to mitigate this bias. We may have reproduced copyrighted content in part by quoting these works, though it was likely not a legal infringement due to the research provision for fair use under US copyright law.

This work did not have any human subjects so it was not necessary to obtain IRB approval.

B Technical Appendices and Supplementary Material

Harms Discussed in ISMIR Ethics Statements		
Harm	**Count**	**Citation**
Copyright	13 (45%)	[5, 7, 9, 23, 30, 33, 38, 46, 49, 50, 56, 57, 65]
Labor	8 (28%)	[30, 37, 48, 49, 50, 56, 58, 62]
General Bias	7 (24%)	[7, 14, 20, 30, 32, 46, 56, 64]
Cultural Appropriation	7 (24%)	[5, 7, 20, 30, 53, 64]
Voice Replication/Impersonation	5 (17%)	[9, 33, 37, 54, 57]
Western Bias	5 (17%)	[1, 8, 9, 58, 63]
Climate	3 (10%)	[7, 26, 30]
Data Scraping	3 (10%)	[7, 22, 58]
Privacy	3 (10%)	[32, 57, 65]
Sustainability	2 (7%)	[7, 37]
Authorship	2 (7%)	[5, 7]
Other	6 (21%)	[5, 16, 20, 48, 49, 62]

Table 1: Harms discussed in the 29 properly utilized ISMIR ethics statements, organized by count of use, percent (of 29), and citations of papers discussing those harms.

Prominent Music AI Works of the 2020s Surveyed in this Work		
Harm/Topic Discussed in Ethics Statement	**Count**	**Citation**
Labor/Economic Impact	7 (54%)	[15, 18, 44, 51, 58, 60, 66]
Legal Acquisition of Data	6 (46%)	[15, 34, 40, 44, 58, 66]
General Bias	6 (46%)	[2, 15, 19, 20, 27, 58]
Western Bias	4 (31%)	[15, 27, 58, 60]
Loss of Agency/Creativity Stifling	4 (31%)	[18, 44, 60, 66]
Voice Replication/Impersonation	2 (15%)	[18, 66]
Copyright	2 (15%)	[51, 60]
Cultural Appropriation	2 (13%)	[2, 27]
No Ethics Statement	12	[10, 12, 17, 21, 24, 28] [29, 39, 43, 42, 41, 45]

Table 2: Harms (or topics) discussed in the 13 ethics/broader impact sections of the 25 prominent Music AI papers surveyed in this work, organized by count of use, percent (of 13), and citations of papers discussing those harms.