2019

Data Usage in MIR: History & Future Recommendations

Wenqin Chen
Smith College

Jessica Keast
Smith College

Jordan Moody
Smith College

Corinne Moriarty
Smith College

Felicia Villalobos
Smith College

See next page for additional authors

Recommended Citation

Wenqin Chen, Jessica Keast, Jordan Moody, Corinne Moriarty, Felicia Villalobos, Virtue Winter, Xueqi Zhang, Xuanqi Lyu, Elizabeth Freeman, Jessie Wang, Sherry Cai, Katherine M. Kinnaird. "Data Usage in MIR: History & Future Recommendations", 20th International Society for Music Information Retrieval Conference, Delft, The Netherlands, 2019.

DATA USAGE IN MIR: HISTORY & FUTURE RECOMMENDATIONS

Wenqin Chen[†] **Jessica Keast**[†] **Jordan Moody**[†] **Corinne Moriarty**[†]
Felicia Villalobos[†] **Virtue Winter**[†] **Xueqi Zhang**[†] **Xuanqi Lyu**
Elizabeth Freeman **Jessie Wang** **Sherry Cai** **Katherine M. Kinnaird**
Smith College, Northampton Massachusetts, USA
kkinnaird@smith.edu

ABSTRACT

The MIR community faces unique challenges in terms of data access, due in large part to country-specific copyright laws. As a result, there is an emerging divide in the MIR research community between labs that have access to music through large companies with abundant funds, and independent labs at smaller institutions who do not have such expansive access. This paper explores how independent researchers have worked to overcome limitations of access to music data without contributing to the crisis of reproducibility. Acknowledging that there is no single solution for every data access problem that smaller labs face, we propose a number of possibilities for how the MIR community can bridge the gap between advancements from large companies and those within academia. As MIR looks towards the next 20 years, democratizing and expanding access to MIR research and music data is critical. Future solutions could include a distributed MIREX system, an API designed for MIR researchers, and community-led advocacy to stakeholders.

1. INTRODUCTION

Since its very conception, the field of Music Information Retrieval (MIR) has struggled with data accessibility, due in part to the nature of music copyright. To deal with this, MIR researchers have developed methods for avoiding or circumventing copyright infringement. Said methods include relying upon public domain and/or Creative Commons music, attainment of certain licenses and/or permissions, use of private in-house data sets, and the access of music data without specifically accessing the audio itself. Each of these methods, however, possess inherent drawbacks such as cost, lack of diverse data, or challenges associated with acquiring required licenses. Public domain and Creative Commons music is often constrained to Western classical pieces; in-house data sets require financial investment; and music data without audio often fail to give a full picture. These drawbacks contribute to an overall disparity of access to data in the MIR community and add to the crisis of reproducibility. In an attempt to strengthen both MIR research and the MIR community as a whole, this paper will propose a number of potential solutions to the drawbacks in existing methods in the hopes that they may serve as a guiding future direction.

The paper is organized as follows. In Section 2, we report on the current concerns and constrictions to data access for MIR researchers. In Section 3, we discuss how datasets from the first ISMIR in 2000 differ from the 10th ISMIR in 2009, and how both compare to the most recent ISMIR in 2018. In Section 4, we outline three proposals for expanding and enhancing access to music data for researchers. Finally, in Section 5, we challenge the community to think urgently about the future and how we can support efforts to improve access to music data.

2. MOTIVATION AND BACKGROUND

Over the last 20 years, society's access to music has evolved. At the first ISMIR conference in 2000, we could only dream of a world with smartphones complete with numerous personalized music streaming applications that make music readily available. Platforms like Spotify and SoundCloud – that now allow for artists to directly share music with the world – had yet to be a pervasive reality.

While the rise of these technologies allows end users to enjoy music more easily and expands MIR research greatly in scope, data commonly used MIR research has not necessarily become more accessible. As with any rapidly growing field, careful attention needs to be paid to legal and scientific concerns with regards to data. In this section, we provide a historical overview of copyright in the United States as it relates to MIR research (though similar statements could be said for other countries). We also discuss the importance of reproducibility for MIR studies.

2.1 History of US Copyright and MIR research

As every MIR researcher is keenly aware, musical recordings are closely protected by copyright laws, which vary by country. In this paper, we treat copyright law in the

† Joint first authors with equal contribution

United States as a case study. The Copyright Act of 1790 marks the beginning of copyright in the US [26]. In 1998, just before the first ISMIR conference, one of the most extensive laws passed was the Digital Millennium Copyright Act (DMCA). This law dealt with issues that arose due to the advent of computers, including allowing data to be copied temporarily during computer maintenance and the ease of broadcast data over the Internet. It also facilitated the implementation of World Intellectual Property Organization (WIPO) Internet treaties, a set of international norms aimed at preventing unauthorized access to and use of creative works on the internet [26]. [1]

As technology has continued to evolve, so too have US laws. In 2018, the Music Modernization Act (MMA) was passed to establish a system of music licensing for digitally distributed music of today [4]. The MMA designates a mechanical licensing collective to oversee a database of musical works that is made publicly available, as well as ensure that artists are paid for their work through the administration of a blanket license. MMA also specifies what songs are in the public domain. With a clearer regulation on music copyright and establishment of a centralized database to track down songwriters, MMA helps reduce unnecessary US copyright related lawsuits while protecting the interests of the music creators. However, given the global nature of the MIR community, being compliant with US laws alone is not enough. As a result, researchers tend to opt for international open source licenses such as Creative Commons (see Section 3.2).

2.2 Reproducibility and Validity of Research

Two vital aspects of scientific research are reproducibility (which encodes reliability) and validity (which often is a proxy for generalizability). These concepts determine how "good" research is – a study that cannot be generalized or trusted is not worth doing.

As MIR tasks grow increasingly complex, it is urgent that the community address the crisis of reproducibility and how it affects the reliability of research. At ISMIR 2014, Raffel et. al. reported that differing evaluation implementations can produce deviations of 9-11% in commonly used metrics across diverse tasks including beat tracking, structural segmentation, and melody extraction [29]. The ability of researchers to verify each other's work is an integral step in the scientific process; without it, consensus on new findings is difficult to reach. That process, however, has been hindered by the use of copyright-protected data in MIR research. Privatized data cannot be legally shared between authors, thus preventing proper re-evaluation for reproducibility [24].

In addition, private and copyright-protected data can be expensive to procure, thus incentivizing and often limiting MIR researchers to using relatively small datasets. This is unfavorable when considering the external validity of a study. It is a general consensus, as Sturm puts eloquently, that "experimental power increases with the number of observations" [35]. Using larger datasets helps control for confounding variables and can sometimes reveal subtle patterns that smaller collections would not pick up on [6]. Smaller datasets carry none of these benefits in addition to having weaker experimental power, greater susceptibility to variation, and revealing only the more obvious patterns. Moreover, limited access to material leads to "restricted and biased subsets," which are difficult to generalize to larger music populations and thus bring into question the external validity of such studies [37].

3. DATA USAGE & EVOLUTION

In this section, we explore data usage and data evolution as well as features of datasets commonly used today. We investigate different classifications of data explored in papers in three different ISMIR conferences. We then discuss the evolution of data creation and examine the limitations of popular datasets.

3.1 Dataset Classification

To illustrate how data usage has evolved over the last 20 years, this paper analyzes datasets used in published papers from ISMIR 2000 [1], 2009 [2], and 2018 [3]. This is done in order to answer three key questions. First, are these datasets diverse enough? Second, how were these datasets created and did the process get easier or harder over time? Third, how were these datasets released and what was their impact?

3.1.1 Genre Classification

To address the first question, we investigate how different music genres have been represented at ISMIR. To do so, we examined the published submissions from the first (in 2000), 10th (in 2009), and most recent (in 2018) ISMIR conferences. Each paper in the respective proceedings was analyzed by three of the authors of this paper, and all music data used for analysis, training, or testing was classified by genre. Although most papers explicitly stated what type of music was used, any disagreements regarding dataset classification were discussed until there was consensus. The results of this survey are displayed in Figure 1. Datasets that used four or fewer genres were classified with each of the genres, those with five or more were given the genre label "various". [2] As many datasets contained multiple genres and many papers referenced multiple datasets, there may be more than one genre representing each paper. Any genre only used once in a given year was categorized as "Other," such as the analysis of opera singing in ISMIR 2018 by Parada-Cabaleiro et.al. [28].

We find that the proceedings in ISMIR 2018 [3] and ISMIR 2009 [2] used data from more music genres, and have a more equal distribution among these genres than those in 2000 [1]. The proportion of pop and rock songs used decreased over time as other genres were included. The proportion of papers that did not use music data or did not

[1] https://www.wipo.int/copyright/en/activities/internet_treaties.html

[2] Datasets with at least one non-Western genre were labeled both "various" and "non-Western." Datasets with only Western genres were given no other genre labels

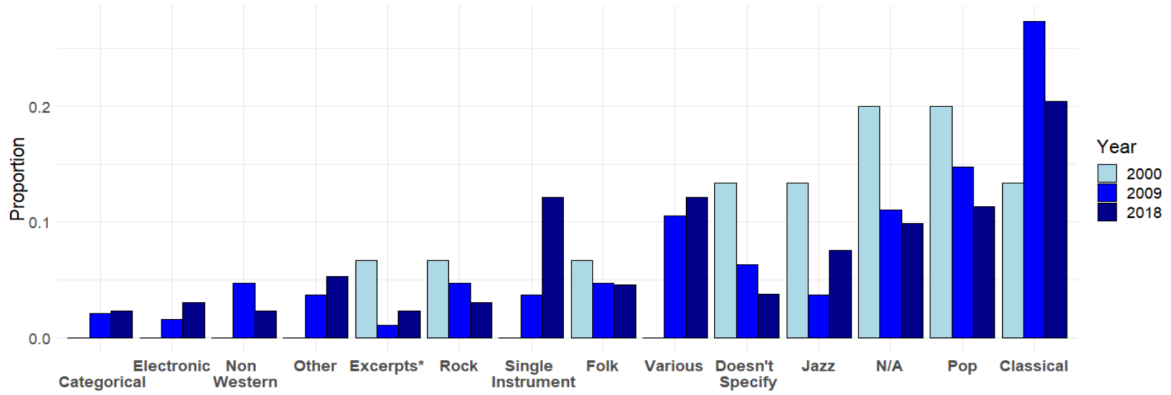

Figure 1: Proportion [3] of the number of papers that use different genres of data from first ISMIR conference in 2000 [1] to the 10th ISMIR in 2009 [2], to the 19th ISMIR in 2018 [3]. "Excerpts*" refers to music excerpts under 3 seconds, and "categorical" refers to music selected for a non-genre category such as mood. The "non-Western" category does not include genres such as J-pop and K-pop, which were classified as solely "pop".

specify the type of music they used have also decreased over time. The latter may be due to an increasing emphasis on reproducibility in MIR, researchers becoming more diligent in documenting their data sources so that others may accurately reproduce their work.

The diversity of MIR research has increased over time. For example, "electronic", "categorical", "non-Western", and "single instrument" music were studied in both 2009 and 2018, even though they had not been in 2000. There has been a substantial increase in papers that consider five or more different genres. However, non-Western genres remain comparatively under-studied, and "classical" continues to be one of the most-studied genres in MIR.

3.1.2 Dataset Collection and Release

To address the second and the third questions, four authors of this work determined the data classifications for datasets used in ISMIR 2018. The categories can be applied to both the dataset creation process and the dataset release process and are summarized in Table 1. [4]

- Creative Commons – CC – includes all datasets released under a Creative Commons license, such as MedleyDB [7]. [5]
- Public Domain – PD – includes all freely available, downloadable, open source datasets that are not licensed under Creative Commons, including datasets released on GitHub; for example, The Million Song Dataset (MSD) [6]. [6]
- Commercial datasets – COM – include those that are owned by for-profit companies such as Spotify and Deezer, and are therefore either not available for public use or available only for purchase.
- In-house datasets – IH – refers to unshareable datasets that are privately owned or contributed via personal connections. For example, some of MedleyDB's data are in-house, though the database is released under CC.
- Library held datasets – LIB – refer to all datasets that are owned and maintained by libraries, or research institutions, including archives dedicated to preserving historical audio records; for example, the University of Iowa Musical Instrument Samples. [7]
- Meta library – MET – refers to aggregates of several existing datasets; for example, the MuMu dataset combines info from MSD and Amazon Reviews [27].
- Permission needed category – PER – refers to any dataset that needs explicit permission to access, but are available for free or at a minimal cost so as to distinguish from the commercial category. An example would be the RWC database [18]. [8]
- Short clips category – SHO – consists of datasets that are made up of short audio samples (usually less than 30 seconds) or recordings of single notes and chords. The NSynth dataset is one example [16]. [9]

3.2 Evolution of dataset creation process: Creative Commons and open source practices

Creative Commons [34] is an international system that creators can use to offer certain usage rights to the public while reserving other rights. This licensing allows content to be distributed within the boundaries of copyright laws, ensuring that creators get credit for their work while allowing for non-commercial distribution and use of the

[3] Proportion out of all genres labels generated for music data used for each paper in that year's ISMIR. The authors found 15 instances of genre usage in ISMIR 2000, 190 in ISMIR 2009, and 132 in 2018.
[4] Note that one dataset can be put into multiple categories.
[5] https://medleydb.weebly.com/downloads.html
[6] http://millionsongdataset.com/
[7] http://theremin.music.uiowa.edu/MIS.html
[8] https://staff.aist.go.jp/m.goto/RWC-MDB/
[9] https://magenta.tensorflow.org/datasets/nsynth

Category	Creation Count	Release Count
CC	10	22
PD	16	32
COM	9	5
IH	18	5
LIB	9	9
MET	9	0
PER	4	9
SHO	12	12
unclear	17	12

Table 1: Classifications of datasets in ISMIR 2018 [3]. Creation count is the number of datasets in each form as gathered by dataset creator, while release count refers to the number of datasets in each form as released for use. There were a number of papers that were not clear on the dataset that they used.

copyrighted material [9, 19]. There are six types of Creative Commons licenses, all applicable worldwide [9, 19].

Before the emergence of Creative Commons in 2002 [8], early MIR research used small and idealized data sets [24]. This is evident in the proceedings of the first ISMIR conference held in 2000, which suffered from a lack of data. Out of the ten papers published in the proceedings, only six described the datasets used. Out of these six, just two papers – [21, 30] – used existing databases, while the other four had to gather data on their own. In comparison with ISMIR 2000, ISMIR 2018 has expanded not just in the scope and complexity of MIR research, but also in the number of datasets referenced, leveraging the resources licensed under Creative Commons or in the public domain.

3.3 Popular Databases Today

Some of the most influential datasets are *Million Song Dataset* (*MSD*), *RWC* and *MedleyDB*. As of June 2019, according to Semantic Scholar, these sets were cited 560, 364, and 124 times respectively. [10] Some CC and OPS datasets such as *Jamendo* are popular, but their influence is difficult to quantify because citations are often not explicitly required. [11] Both *MedleyDB* and *Jamendo* rely on Creative Commons for copyright-free distribution of music files [7, 20].

In ISMIR 2018 alone, the *MSD* was used in 11 different papers, making it one of the most frequently cited datasets in that conference. Released in 2011, the *MSD* is a freely available collection of derived Echo Nest features and metadata of one million contemporary Western commercial songs [6]. The dataset can be used along with `7digital` to fetch short samples of songs within certain limitations for free [6].

This lack of audio is not unusual; our investigations (described in Section 3.1) found that about one-fifth of all datasets used in ISMIR 2018 do not contain any audio files. Research on this kind of data is limited to the provided features (by Echo Nest in the case of *MSD*), and does not allow for the kind of in-depth investigation that audio data provides. However, the model of collaborating with free, community-driven datasets as well as commercial companies has served as a good solution for expanding the scale of MIR datasets without violating copyright law.

Due to copyright limitations, many MIR researchers also choose to use music databases that consist exclusively of public domain works, Creative Commons works, or those with limited features of the raw audio. The *Classical Music Archive*, for instance, is a large repository of classical music with each track existing in public domain [5]. The pervasiveness of classical music (and dirth of other genres) in the public domain leads to research being greatly skewed towards the classical genre with other genres being explored much less. Although this research remains important, it is worth pointing out that the vast majority of music produced today is not classical, and is not encompassed by these studies.

Other datasets avoid copyright infringement by omitting material subject to copyright. For instance, *AcousticBrainz* is a publicly available database that is entirely composed of features extracted from songs, rather than the songs themselves [5]. *AcousticBrainz* extracts features on the song level, accumulating the values researchers might want to use and associating them with songs without storing any audio [5]. Similarly, in 2006, *OMEN* was proposed as a system of feature extraction that would take place in libraries [22]. The proposal indicates a method of communication between researchers and libraries that would allow researchers to request specific features from a specific song that would then be extracted by librarians. Frameworks such as *OMEN* would help researchers circumvent copyright and the cost of in-house datasets while protecting content created by artists through the separation of music from its features to create more available music data.

3.4 Many Datasets; Yet Limited Data & Data Access

Despite the numerous cited datasets, an issue repeatedly raised within MIR is the existence of bias [12] in data commonly used in the field. Bias is frequently inevitable due to lack of resources. For example, the structural segmentation task in MIREX 2010 used two datasets with 397 songs in total almost exclusively biased towards Western popular music [15]. This was partially because the data were donated by a few universities and were therefore limited in quantity and diversity. A new annotated dataset with over 1,300 songs covering pop, jazz, classical and world music then became available in 2011 [15], and has been used since in the structural segmentation task of MIREX.

Community and partnership-driven data initiatives have provided working data access to further MIR research. However each member of the MIR community faces

[10] Accessed on June 23, 2019

[11] https://www.jamendo.com/

[12] Bias can exist in terms of geographic location, musical instruments, music genres, musical styles and musical fundamentals such as key, harmony, tempo and rhythm.

unique hindrances to accessing data. *MSD* and Creative Commons currently help researchers bypass these challenges with innovative solutions, but do not support a data access framework for the MIR community as a whole.

4. PROPOSALS

As concerns about access and reproducibilty mount, we offer proposals to increase data access that draw upon resources existing within the MIR community. These suggestions seek to address the crisis of reproducibility while working within current copyright law.

The three proposals below tackle these issues from three different parts of the MIR community. The first calls upon our academic partners to build a distributed MIREX. The second asks our industrial partners to leverage their existing infrastructures. The third asks the whole community to expand our support for dataset creators. Noting that MIR researchers around the world fall under different copyright laws given their location, some of these proposals focus on actions that could be taken in the United States (to the benefit of the worldwide community), while others provide potential global solutions.

4.1 Academic Proposal: Distributed MIREX

Our first proposal is a new, decentralized, distributed Music Information Retrieval Exchange (MIREX) system that would build on the success of the existing MIREX infrastructure, while leveraging previously untapped resources to support the development of such a system. Drawing on the strengths of the proposals in [11–13], this distributed MIREX seeks to be a middle ground between small, in-house datasets and large industrial ones – allowing researchers to test algorithms without violating copyright or relying solely on non-copyrighted music.

This proposed system would distribute the responsibilities and challenges associated with running MIREX among three institutions, by creating two new MIREX sites in addition to the current one run by IMIRSEL at University of Illinois Urbana-Champaign (UIUC). Though the startup cost of a new MIREX site is substantial (data, servers, and other infrastructure needs), creating several MIREX locations is a long-term investment in MIR research and future researchers. More MIREX locations will increase the community's opportunities to access music datasets and allow for additional comparisons between algorithms.

4.1.1 Current MIREX

MIREX was created by IMIRSEL as a method to make comparisons between algorithms. The system's process of submitting, running, and returning results ensures that the music data being used is in accordance with US copyright laws [10]. Though MIREX has been a successful tool for MIR research, it presents challenges of accessibility and efficiency. In response to these challenges, several MIR papers have called on the community for help and have proposed a distribution of MIREX responsibilities including: a web service system to give some algorithm execution responsibility to the user [13], a distributed task management [11], a breakdown of the MIR process across labs participating in the NEMA framework with inter-lab communication [12], and a proposal for a distributed computation model that leverages open source methods [23].

4.1.2 Logistics of Distributed MIREX

The first step in creating a distributed MIREX is to identify the additional two sites. We propose that the three sites would not operate in a coordinated fashion, as suggested in [11, 12], other than ensuring that the timings of the runs do not overlap. Undergraduate-focused institutions – common in the US – are ideal for selection as new MIREX locations, as these schools tend to have existing systems for supporting undergraduate student research and student work-study programs.

Within the MIR community, MegsRadio, under the direction of Douglas Turbull, is an example of an undergraduate student driven MIR project that thrived at Ithaca College [36]. MegsRadio ran throughout the year with students addressing technical issues as well as working on user-facing elements of their work [36]. Student involvement similar to that in MegsRadio is what we envision for the proposed distributed MIREX. Additionally, research has suggested that students who do undergraduate research in STEM may be more likely to pursue a graduate/professional degree in a STEM field [14]. It stands to reason that these findings could be extended to MIR, especially the field's technical aspects. As such, intentionally involving more undergraduates in MIR research will help broaden the field and provide some of the labor required to maintain a distributed MIREX system.

In addition to the potential of undergraduate workers, our distributed MIREX's ability to run algorithms on different datasets multiple times per year could allow researchers the the chance to test their work more than once a year. Furthermore, this system could lessen the individual infrastructure-related responsibility of handling the "interim data explosion" problem [12] caused by specific algorithms. More locations could allow for more storage of features created and generated by algorithms, which the current MIREX discards due to storage limitations [12]. The task of maintaining not only the data itself, but any features created by algorithms, would be distributed to various MIREX sites – thus addressing the infrastructure capacity issues in the current model. The current system entails engaging in several hours of intensive back and forth communication with participants about their algorithms as described in [11]. Instead, this distributed MIREX could increase access to more diverse and quality datasets over time as well as foster growth in the MIR community. Additionally, the creation of the distributed MIREX would compliment open source systems like those proposed in [23] at ISMIR 2016.

4.2 Industrial Proposal: Researcher API License

As the MIR field has grown, so have industrial research groups and public-facing technologies. To complement

our academic proposal, we advocate developing a researcher API license that leverages existing industrial infrastructure. This proposed new license would be designed with the workflow of MIR research practices in mind and would allow for actions that are completely disjointed from current developer licenses.

Developer licenses are designed with software developers in mind and possess inherent restrictions that make them unsuited for MIR research. For example, a streaming phone app that allows for streaming from a different platform would be created under a developer license. But actions such as feature extraction, harmonic-percussive separation, and sampling are in violation of most developer licenses (for example, see Section IV.1.a of [33]). These guidelines leave very little room for MIR research of any kind, and similar language exists on other APIs' terms of use, including Soundcloud's API terms of use that prohibits the using of any data from their platform (see [32]).

We propose that companies with existing developer licenses and API infrastructure (such as Spotify and SoundCloud) create a new license that explicitly allows for common MIR practices such as feature extraction and explicitly prohibits using algorithm outputs for commercial or listening purposes, but limits casual listening. To create appropriately narrow language, companies could consult with IMIRSEL who forged similar agreements to build the MIREX [10]. Such a license addresses concerns about data biased towards popular music, as smaller artists can directly share their work on these streaming platforms.

An immediate and real concern with such a license is the potential for abuse. To help address this concern, we propose the creation of an ethics training module similar to those taken by researchers conducting work using human and animal subjects (colloquially referred to as IRB courses). In this ethics training module, MIR researchers would be educated on the ethical, cultural, and financial issues at play in using and misusing music data. MIR researchers would also have to complete various assessments that certify their understanding of these nuanced issues.

This proposal has the potential for global reach as many companies already operate in several countries, and seeks to provide access to more diverse data.

4.3 Community Proposal: Usage and Advocacy

In addition to creating a decentralized MIREX system and an academic API license, the MIR community must also continue to support the current creators and maintainers of research datasets, not only through their use but through advocacy of laws that make them possible in the first place. The passing of laws such as the Music Modernization Act in the US (Section 2.1) affect the MIR community, but the community had no influence on the creation or implementation of this policy.

Since ISMIR is affected by these policy changes, it is important for the society to know about the current policies. It would be beneficial to be involved in the politics behind these laws to ensure that the voices of researchers are heard. Many research societies have created frameworks for government communication. The American Association for the Advancement of Science (AAAS) has the Center for Science Diplomacy, which connects science to policy and allows the strengthening of connections between diplomatic and science communities [17]. The American Mathematical Society (AMS) is a part of the Coalition for National Science Funding and the Task Force on American Innovation, both of which are alliances of professional, academic, and scientific organizations that advocate for issues that affect the members and their communities [31]. These alliances and committees ensure that these fields of research are not negatively affected by policy changes and that the voices of the societies are heard.

As stated on the ISMIR website, one of the purposes of the ISMIR Organization is "to cooperate with representatives of other organizations and disciplines toward the furtherance of music information retrieval" [25]. To achieve this goal, it is necessary to cooperate not only with other research societies but also with the government bodies that create the infrastructures that MIR researchers work under. The creation of an ad-hoc committee or some other system focused on current music policy changes and advocating for music research in policy would benefit the MIR community. Allying with other music research organizations such as the American Musicological Society or the International Musicological Society could also help facilitate change in music policy for the betterment of research. These systems would benefit both the society and the growth of the MIR field as a whole.

5. CONCLUSION: TOWARDS A STRONGER MIR COMMUNITY

We have proposed three possibilities for expanding researcher access to music data. While progress on any of these three proposals will better the field, simultaneous efforts will lead to a collective impact that is greater than the sum of each proposal's individual impacts. For example, as new MIREX locations begin, industrial researchers can consult on how to design and build appropriate infrastructure given the school's constraint. Similarly, in the US model, undergraduate students working in MIR research groups would make excellent testers for early versions of companys' researcher API licenses. Finally, with the creation of an ad-hoc committee (supported by the ISMIR board) representing the interests of MIR researchers to policy-makers and vice versa, the whole MIR community will be better informed about current laws concerning music. This increased awareness will help to continually improve both academic and industrial research efforts.

Unequal access to music data has led to field-wide issues including a crisis of reproducibility and concerns about access. Despite this, future directions and solutions do exist. By implementing the multi-pronged approach outlined in this paper, MIR research can adapt to overcome these obstacles and avoid the looming schism between corporate and independent labs. Moreover, the methods we have proposed will foster cooperation within the field and could lead to greater diversity within MIR.

6. ACKNOWLEDGEMENTS

The last author is the Clare Boothe Luce Assistant Professor of Computer Science and Statistical and Data Science at Smith College and as such, is supported by Henry Luce Foundation's Clare Boothe Luce Program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Luce Foundation.

7. REFERENCES

[1] ISMIR 2000: International Symposium on Music Information Retrieval. http://ismir2000.ismir.net/papers.html, 2000. [Online; accessed 3-April-2019].

[2] ISMIR 2009: 10th International Society for Music Information Retrieval conference. https://ismir2009.ismir.net/program.html, 2009. [Online; accessed 20-June-2019].

[3] ISMIR 2018: 19th International Society for Music Information Retrieval conference. http://ismir2018.ircam.fr/pages/events-main-program.html, 2018. [Online; accessed 3-April-2019].

[4] 115th Congress. H.R.5447 - Music Modernization Act. https://www.congress.gov/bill/115th-congress/house-bill/5447/text, 2018. [Online; accessed 10-April-2019].

[5] Y. Bayle, M. Robine, and P. Hanna. SATIN: a persistent musical database for Music Information Retrieval and a supporting deep learning experiment on song instrumental classification. *Multimedia Tools and Applications*, 78(3):2703–2718, 2019.

[6] T. Bertin-Mahieux, D. P. W. Ellis, B. Whitman, and P. Lamere. The Million Song Dataset. In *Proc. of* 10^{th} *ISMIR Conference*, pages 591–596, 2011.

[7] R. M. Bittner, J. Salamon, M. Tierney, M. Mauch, C. Cannam, and J. P. Bello. MedleyDB: A multitrack dataset for annotation-intensive MIR research. In *Proc. of* 15^{th} *ISMIR Conference*, pages 155–160, 2014.

[8] Creative Commons. The story of Creative Commons. https://certificates.creativecommons.org/cccertedu/chapter/1-1-the-story-of-creative-commons/. [Online; accessed 28-June-2019].

[9] Creative Commons. About the licenses. https://creativecommons.org/licenses/, 2018. [Online; accessed 10-April-2019].

[10] J. S. Downie, J. Futrelle, and D. K. Tcheng. The international Music Information Retrieval systems evaluation laboratory: Governance, access and security. In *Proc. of* 5^{th} *ISMIR Conference*, 2004.

[11] J. S. Downie, X. Hu, J. H. Lee, K. Choi, S. J. Cunningham, and Y. Hao. Ten years of MIREX: Reflections, challenges and opportunities. *Proc. of* 15^{th} *ISMIR Conference*, pages 27–31, 2014.

[12] J.S. Downie. The Music Information Retrieval Evaluation Exchange (2005-2007): A window into Music Information Retrieval research. *Acoustical Science and Technology*, 29(4):247–255, 2008.

[13] J.S. Downie, A. Ehmann, and J. Lee. The Music Information Retrieval Evaluation eXchange (MIREX): Community-led formal evaluations. *Digital Humanities (Conference)*, 29(4):239–241, 2008.

[14] M. K. Eagan Jr., S. Hurtado, M. J. Chang, G. A. Garcia, F. A. Herrera, and J. C. Garibay. Making a difference in science education: The impact of undergraduate research programs. *American Educational Research Journal*, 50(4):683–713, 2013. PMID: 25190821.

[15] A. F. Ehmann, M. Bay, J. S. Downie, I. Fujinaga, and D. De Roure. Music structure segmentation algorithm evaluation: Expanding on MIREX 2010 analyses and datasets. In *Proc. of* 12^{th} *ISMIR Conference*, pages 561–566, 2011.

[16] J. Engel, C. Resnick, A. Roberts, S. Dieleman, M. Norouzi, D. Eck, and K. Simonyan. Neural audio synthesis of musical notes with wavenet autoencoders. In *Proc. of the 34th International Conference on Machine Learning - Vol. 70*, ICML'17, pages 1068–1077. JMLR.org, 2017.

[17] The American Association for the Advancement of Science. Science diplomacy. https://www.aaas.org/focus-areas/science-diplomacy, 2019. [Online; accessed 11-April-2019].

[18] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka. RWC music database: Popular, classical and jazz music databases. In *Proc. of* 3^{rd} *ISMIR Conference*, pages 287–288, 2002.

[19] C. Green. Open licensing and open education licensing policy. In *Open: The Philosophy and Practices that are Revolutionizing Education and Science*, pages 29–42. Ubiquity Press, 2017.

[20] Jamendo. Jamendo music. https://www.jamendo.com/, 2019. [Online; accessed 8-April-2019].

[21] P. Jeremy. A comparison of language modeling and probabilistic text information retrieval approaches to monophonic music retrieval. In *Proc. of* 1^{st} *ISMIR Conference*, 2000.

[22] D. McEnnis, C. McKay, and I. Fujinaga. Overview of OMEN. In *Proc. of* 7^{th} *ISMIR Conference*, pages 7–12, 2006.

[23] B. McFee, E.J. Humphrey, and J. Urbano. A plan for sustainable MIR evaluation. In *17th International Society for Music Information Retrieval Conference*, ISMIR, 2016.

[24] B. McFee, J. W. Kim, M. Cartwright, J. Salamon, R. M. Bittner, and J. P. Bello. Open-source practices for music signal processing research: Recommendations for transparent, sustainable, and reproducible audio research. *IEEE Signal Processing Magazine*, 36(1):128–137, Jan. 2019.

[25] The International Society of Music Information Retrieval. About the society. `https://www.ismir.net/society.php`, 2017. [Online; accessed 11-April-2019].

[26] Association of Research Libraries. Copyright timeline: History of copyright in the United States. `https://www.arl.org/focus-areas/copyright-ip/2486-copyright-timeline`, 2019, continually updated. [Online; accessed 08-April-2019].

[27] S. Oramas, O. Nieto, F. Barbieri, and X. Serra. Multi-label music genre classification from audio, text and images using deep features. In *Proc. of* 18^{th} *ISMIR Conference*, pages 23–30.

[28] E. Parada-Cabaleiro, M. Schmitt, A. Batliner, S. Hantke, G. Costantini, K. Scherer, and B. W. Schuller. Identifying emotions in opera singing: Implications of adverse acoustic conditions. In *Proc. of* 19^{th} *ISMIR Conference*, pages 376–382, 2018.

[29] C. Raffel, B. McFee, E. J. Humphrey, J. Salamon, O. Nieto, D. Liang, and D. P. W. Ellis. `mir_eval`: a transparent implementation of common MIR metrics. *Proc. of* 15^{th} *ISMIR Conference*, 2014.

[30] C. G. Sayeed, D. Michael, D. Tim, F. Ichiro, and H. Brian. Optical music recognition system within a large-scale digitization project. In *Proc. of* 1^{st} *ISMIR Conference*, 2000.

[31] The American Mathematical Society. Office of government relations collaborations and position statements. `https://www.ams.org/government/dc-collaborations`, 2019. [Online; accessed 11-April-2019].

[32] Soundcloud. API terms of use. `https://developers.soundcloud.com/docs/api/terms-of-use`, 2016. [Online; accessed 5-April-2019].

[33] Spotify. Spotify developer terms of service. `https://developer.spotify.com/terms/`, 2018. [Online; accessed 3-April-2019].

[34] E. Steuer. Response to ASCAP's deceptive claims. `https://creativecommons.org/2010/06/30/response-to-ascaps-deceptive-claims/`, Jun. 2010. [Online; accessed 28-June-2019].

[35] B. L. Sturm. Revisiting priorities: Improving MIR evaluation practices. In *Proc. of* 17^{th} *ISMIR Conference*, pages 488–494.

[36] D. Turnbull, C. Perez, D. Dorsey, J. Burger, J. Menduni, E. Piech D. Akimchuk, A. Python, and T. Joachims. Developing a professional-grade personalized radio app in an academic setting. *Late breaking at* 17^{th} *ISMIR Conference*, 2017.

[37] J. Urbano, M. Schedl, and X. Serra. Evaluation in Music Information Retrieval. *Journal of Intelligent Information Systems*, 41:345–369.