



Amplifying Artists' Voices: Item Provider Perspectives on Influence and Fairness of Music Streaming Platforms

Karlijn Dinnissen

k.dinnissen@uu.nl

Utrecht University

Utrecht, The Netherlands

Christine Bauer

c.bauer@uu.nl

Utrecht University

Utrecht, The Netherlands

ABSTRACT

The majority of music consumption nowadays takes place on music streaming platforms. Whichever artists, albums, or songs are exposed to consumers on these platforms therefore greatly influences what music is ultimately consumed. As a result, the impact of these platforms on artists—their main item providers—is considerable. The recommender systems at the core of streaming platforms, though, have traditionally been developed focusing on end consumer objectives. Only recently, researchers have started to include item provider objectives, though rarely through reaching out to item providers directly. By omitting this important stakeholder's point of view, we risk not understanding what artists value most, and might miss first-hand ideas on how to improve streaming platforms and recommender systems. Therefore, we conducted semi-structured interviews to capture the artists' view. Specifically, we explore artists' considerations regarding fairness, transparency, and diversity in music recommender systems, and the role artists envision for streaming platforms regarding those topics. We identify some topics with a clear consensus among artists, such as desiring more control over which music is recommended to whom, and expecting streaming platforms to actively increase music diversity in recommendations. In contrast, artists' opinions differ on whether platforms should actively intervene in recommender systems to, e.g., increase localization or gender balance. Further, we observe that artists often take user preferences into account and even suggest new platform functionality to benefit both users and item providers. We encourage utilizing these insights when designing and evaluating music streaming platforms and recommender systems.

CCS CONCEPTS

• **Information systems** → **Recommender systems**; • **Human-centered computing** → **Empirical studies in collaborative and social computing**;

KEYWORDS

fairness, music recommender systems, human-centered computing, interviews, item provider



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs International 4.0 License.

UMAP '23, June 26–29, 2023, Limassol, Cyprus

© 2023 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9932-6/23/06.

<https://doi.org/10.1145/3565472.3592960>

ACM Reference Format:

Karlijn Dinnissen and Christine Bauer. 2023. Amplifying Artists' Voices: Item Provider Perspectives on Influence and Fairness of Music Streaming Platforms. In *UMAP '23: Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP '23)*, June 26–29, 2023, Limassol, Cyprus. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3565472.3592960>

1 INTRODUCTION

Music streaming platforms have become one of the major resources for music consumption [44, 61]. With a vast amount of music tracks available on these platforms¹, music recommendation has not only become a commodity to help music listeners navigate this enormous offer, but even a necessity. The most common use for music recommender systems (MRS) is to provide a user with the most suited tracks or artists in the right context [55]. Users then pay attention to what they are exposed to [10]. As a result, what is presented on music streaming platforms—and what is promoted through the algorithmic recommendations in particular—strongly influences what users will consume there. This, in turn, impacts music artists, whose contributions make up the core value of the platforms: artists are their main item providers [2]. Eventually, exposure and attention to the items shape the music streaming ecosystem at large [23].

Due to their prevalence and impact in many domains, there is a general need to make online platforms fair environments [13]. A big challenge in this realm is to define what fairness connotes and entails, as this may vary across stakeholder groups [59]. What is more, when developing, optimizing, and evaluating a recommender system (RS), it is important to specifically define and address the goal that should be achieved with the very system in the very context [33, 71]. This may require looking into domain and application specifics, as findings from one context do not necessarily generalize to other contexts [32, 33]. This also holds true for research on and development of fair RS. In our work, we zoom in on the music domain; we delve into the specific context of music streaming platforms with the ultimate goal of encouraging increasing artist fairness in MRS.

In the music domain, several works address RS fairness. This can be done from different stakeholders' perspectives (for a recent overview, see Dinnissen and Bauer [11]). For instance, fairness has been addressed from the user perspective (e.g., [4, 14, 38]), the artist perspective (e.g., [17, 22, 49]), and the perspectives of multiple stakeholders (e.g., [44, 49]). Despite the existence of such different perspectives, the majority of works to date address user fairness [11]. In comparison, research addressing item provider

¹80–90 million tracks as of 2022, cf. <https://newsroom.spotify.com/company-info/> and <https://www.apple.com/apple-music/>

fairness is scarce, with the exception of, e.g., Ferraro et al. [22] addressing the gender imbalance among recommended artists, and Knees et al. [36] investigating biases across record labels.

In addition, researchers of RS in general have rarely reached out to item providers to understand what they consider fair, and especially how a RS should work to be considered a fair one. The music domain is no exception, with very limited research (e.g., [23]) directly reaching out to item providers. Therefore, our research builds on and extends this previous work. While Ferraro et al.'s [23] study is limited to a sample of 9 artists, all from Spanish-speaking countries, we address the topic in a different cultural setting: the Netherlands. This country moderately to strongly differs in all six of Hofstede's cultural dimensions [29], compared to the countries from which the Spanish-speaking artists originate [30]. Studies in different cultural contexts provide insights far beyond reproducibility as, e.g., recently shown in the context of music information retrieval [37]. We also add to previous work by considering a larger sample here, allowing us to present the views of more different types of artists. Lastly, the COVID-19 pandemic has changed the music landscape [34, 58], but, again, no work has reached out yet to artists to understand the impact of these changes on their fairness perspectives regarding MRS. In our work, we therefore explore the artist perspective on fairness of music streaming platforms and in particular their MRS, guided by the following research questions:

- **RQ1:** What do artists consider to be fair in music streaming platforms and embedded recommender systems?
- **RQ2:** Which role do artists envision for music streaming platforms with regard to fairness, diversity, and transparency of their recommender systems?

The contribution of our work is three-fold. First, for a set of fairness aspects, we confirm the findings of Ferraro et al. [23], yet in another cultural context. Second, our study contributes valuable new insights concerning what artists consider fair: additional topics are introduced, and on some fairness aspects our sample's responses contradicted or were more nuanced than previous studies. Third, our work offers concrete ideas for new user interface (UI) functionality that could increase fairness. With the insights our research adds, we contribute to a *better understanding of fairness in MRS from the artist perspective*, and provide an *informed basis for the design of fair MRS*.

The remainder of this work is structured as follows: In Section 2, we present the conceptual basis and discuss related work. Then, in Section 3, we detail the methods, including the interview procedure and participant description. Section 4 constitutes the core part of this paper, where we present insights from the interviews. We discuss our results in Section 5, and conclude with a summary of contributions, limitations, and future research directions in Section 6.

2 RELATED WORK

Here, we first show that RS involve multiple stakeholders. We outline fairness conceptualization work, demonstrating that fairness is multifaceted and embraces issues like popularity bias, diversity, and the cold start problem. The discussion shows that it is an open issue to define 'what is fair', making it important to involve and understand all stakeholders.

Fairness and stakeholders of music recommender systems. The advent of music streaming platforms has changed the music landscape [5, 67], which has strongly affected artists [12]. Therein, the RS that are integrated into these platforms play an especially important role [48]. Many factors make these RS prone to biases, resulting in unfair outcomes [13] for end consumers (users) [14, 62] and item providers [3, 15] alike. This holds true in the music domain (e.g., [4, 38] for user perspective, [17, 22] for artist perspective). However, even though RS may affect both user and item provider fairness, most research focuses on recommendation value for users [32, 45]. This phenomenon also occurs in general research on fairness in information access systems [13], and in MRS research specifically [11]. As an exception, a recent survey on RS fairness discusses more publications on item provider fairness than user fairness [69]. In this work, we add new insights about the item provider perspective by putting artists in the loop.

Conceptualization of fairness. There is a multitude of definitions of (algorithmic) fairness; for an overview, see Hutchinson and Mitchell [31]. A commonly adopted approach is to distinguish two main categories in measuring fairness: *individual fairness* and *group fairness*. Individual fairness requires that individuals are treated in a similar way. Group fairness compares the outcome across groups and seeks some type of equivalence between groups of individuals that share a characteristic. Hereby, the challenge is that we may define individuals as similar or dissimilar according to various characteristics. For instance, anti-discrimination regulations clearly define so-called 'protected groups' according to characteristics such as gender, age, ethnicity, and nationality, and explicitly prohibit unfair treatment for those protected groups [1, 7, 18, Art. 21]. We note that group fairness compares the outcome; thus, while it might be useful to include characteristics of (protected) groups in user modeling to allow for a better adapted service, this must not result in outcomes that discriminate against a protected group.

However, unfairness may also occur on other grounds than characteristics protected by law. For instance, many recommendation approaches are prone to popularity bias [16], which may, e.g., increase the gap between highly popular artists and (in comparison) less popular ones. This issue is particularly accentuated for new artists, who first need to reach a decent amount of item ratings before a MRS can consider their items in the recommendations (new-item cold start problem [56]). Popularity bias may, in turn, also result in fairness issues from the user perspective, as users with preferences aligned with the mainstream tend to get better recommender results compared to users with preferences beyond what is generally popular [38]. Further, fairness issues may arise with regard to artists of different career lengths [49], different genres [49], or across artist types (e.g., solo artists, bands) [21]. In addition, we have to consider *intersectionality* [8], as sometimes unfairness may only be revealed by looking at combinations of characteristics.

In the conceptualization of fairness, domain-specifics should be considered. For instance, viewpoint diversification is relevant in news recommendation [65] but not in the fashion domain. Fairness across genres might be relevant in music, but not in the health domain where 'genre' is irrelevant. Further, gender imbalance is a highly topical subject in the music sector (e.g., [28, 68, 70]) and may thus be particularly relevant there. In domains related to cultural

content, research on the broader social role also needs to center culture as a core concept and challenge [20].

It is important to note that, though fairness and diversity are different [50], there is a link between these concepts [46]. While diversity is more about the variety in the representation of individuals², fairness is about fair outcomes (i.e., not favoring some individuals based on certain characteristics) [50]. Pitoura [50] points out: “A diverse output is not necessarily a fair one, and a fair output is not necessarily a diverse one.” For example, genders might be equally represented in a set of people, which may be considered diverse, but gender groups may still face different outcomes (e.g., different pay, one gender group clustered at the end of a ranked list), which may be considered unfair. Nevertheless, diversity may often be a means to nurture fairness [6, 35].

Similarly, transparency is often discussed in a fairness context. Essentially, it is a prerequisite that helps assess whether a RS is fair. In this context, Sonboli et al. [60] point out that claiming a RS is fair is not enough; people need to understand the encoded fairness objectives. To this end, transparency helps to perceive a fair system indeed as fair. We will explore artists' considerations regarding fairness of MRS, and specifically capture which aspects they consider most relevant on this topic.

Stakeholder involvement. Notably, research inquiring stakeholders directly about their desires and needs concerning RS fairness is scarce [13], though some first research surfaces in this direction. For instance, some works take the user perspective and explore users' opinions on what fair treatment connotes in the context of algorithmic recommenders, and which RS features and capabilities are desired in this respect [26, 60]. With a large-scale survey, Helberger et al. [27] investigate why and under which conditions people accept algorithmic decision-making as fair. Ferraro et al. [23] take the item provider perspective and study music artists' perceptions and demands concerning fairness in the context of MRS. While not addressing fairness specifically, the work by Siles et al. [57] demonstrates that it is impactful to reach out to artists to better understand the impact of music streaming platforms—and playlists in particular—on artists. Overall, though, little is to date known about what users or item providers expect from a RS with respect to fairness [13]. In this work, we conduct interviews with artists to understand their perspective on the impact of music streaming platforms, and also explore their ideas for potential improvement to eventually help inform the design of fairer MRS.

3 METHODS

To understand artists' perspectives on the current music streaming platforms and their embedded MRS, we conducted 14 semi-structured interviews in the Netherlands from January through March 2022. As the interviews took place during the COVID-19 pandemic, all interviews except one were held online. Here, we describe the participant sample (Subsection 3.1), the interview process and protocol (Subsection 3.2), and the analytical process (Subsection 3.3). The research methods were reviewed and approved by the Utrecht University Science-Geosciences Ethics Review Board.

²Mitchell et al. [46] emphasize that diversity denotes variety in terms of social attributes such as sociopolitical power differentials.

3.1 Participants

We reached out to potential interviewees via their management, social media accounts, and through personal networks (and their respective networks). We paid particular attention to diversity in age, gender, audience reach (i.e., international, national, local)³, genre, lyrics language, and record label affiliation⁴. This resulted in 14 interviews with currently active music artists living in the Netherlands. For music groups consisting of several members, we offered the possibility for two members to be interviewed together. Therefore, 3 interviews were held with two artists simultaneously, and 11 interviews with one artist. Table 1 provides an overview of all participants, with self-reported personal characteristics.

As the purpose was to provide an in-depth and illustrative understanding of Dutch artists' perspectives, we continued reaching out to new participants until we reached a high level of thematic saturation [25]. Considering our sample diversity, and our sample size which corresponds to common research practice [9, 47], the final sample proved to match our objectives. Yet, we note that we do not aim to offer a view that is representative for all artists, but rather to expand existing insights that focused solely on artists from Spain and Latin America [23, 57].

3.2 Interview Protocol

All materials were prepared in Dutch with Ferraro et al. [23] as a basis. These consist of a letter of invitation, an informed consent form, a metadata questionnaire, presentation slides, and interview questions.⁵

Consent form. We discussed the privacy procedures put in place to protect participants' identities. We emphasized that interview audio and transcripts are kept strictly confidential, and that results are only reported in aggregated form or pseudonymized. Hereby, we aimed to create a safe atmosphere in which artists would feel free to share concerns or issues regarding music streaming platforms or the music industry, without fearing potential negative repercussions.

Metadata questionnaire. Participants were asked to fill out general demographics and music background information (Table 1). Where possible, questions contained categories as opposed to free text fields, to prevent responses that are specific enough to identify an artist. Participants were also free to skip any question.

Presentation. We developed slides containing a ~12-minute introduction to MRS and their integration into music streaming platforms. With this, we aimed to give participants a foundational understanding before their interview.

Interview questions. For our semi-structured interviews, we used a set of neutrally phrased guiding questions, and optional follow-up questions to encourage elaboration. Table 2 provides an overview of the questions; exact phrasing can be found in our materials' package. They were encouraged to think of music recommendation in several shapes and forms, and how any issues could be improved. The questions referred to music streaming platforms in general, not singling out any platform in particular. As we build

³Note that a national audience reach does not necessarily mean that the artist is only active within one country, but rather that they consider their reach, popularity, and social influence to be on a national level.

⁴Major label refers to Sony Music, Universal Music Group, or Warner Music Group. Indie refers to labels that are independent of any major label.

⁵All materials can be accessed at <https://doi.org/10.5281/zenodo.7883507>.

Table 1: Participant information (self-reported).

Code	Age	Gender	Audience reach	Genre	Lyrics language ^a	Label affiliation
P1	26–35	Male	Local	Hip-Hop	nl	Self-published
P2	26–35	Male	National	Rock/Pop	en	Owns rights, label publishes
P3	26–35	Male	Local	Rock, Punk, Metal	en, instrumental	Self-published
P4a	26–35	Male	(P4a) Local	(P4a) Hardcore, Rock, Blues	(P4a) nl, en, instrumental	(P4a) Indie label + self-published
P4b			(P4b) National/Local	(P4b) Indie, Metal, Cl. Rock	(P4b) en, instrumental	(P4b) Self-published
P5	26–35	Male	International	Dance	en, instrumental	Major label
P6	18–25	Non-binary	Local	Pop	nl, en, instrumental	Self-published
P7a	46–55	(P7a) Female	National	Alt. Pop	en	Self-owned label
P7b		(P7b) Male				
P8a	56–65	Female	N/A	Folk, World	Self-invented language	Self-published
P8b						
P9	18–25	Non-binary	Local	Rock, Pop, Folk	en	Self-published
P10	26–35	Male	Local	Neoclassical	instrumental	Major label + self-published
P11	36–45	Female	Local	80's Alt. Synthpop	en	Self-published
P12	18–25	Female	Local	Metal	en, instrumental	Indie label
P13	26–35	Female	(Inter)national	Indie-pop Alt.	en, instrumental	Indie label
P14	36–45	Male	National	Many	nl, en, instrumental	Major label

^aWe use ISO 639-1 2-letter codes to refer to languages; cf. <https://www.iso.org/iso-639-language-codes.html>

on Ferraro et al. [23], questions were kept mostly the same, with some changes to order and phrasing. We added Q11 to address the impact of the COVID-19 pandemic on their career and view on music streaming platforms, as the music landscape has changed since 2020 [34, 58] and Ferraro et al. [23] conducted the interviews before the pandemic.

Process. Participants were invited to read and sign the conditions for informed consent before their interview. After this, they were asked to fill out the metadata questionnaire, and listen to the interviewer's MRS presentation. Then, the interview recording started. The remaining time (on average 51 minutes and 45 seconds) was dedicated to the questions as outlined above, with the longest interview lasting 60 minutes and 55 seconds, the shortest 43 minutes and 15 seconds.

3.3 Data Processing and Qualitative Content Analysis

The audio of each interview was recorded, transcribed, and pseudonymized, with the pseudonymized transcripts counting 113,357 words in total. A Qualitative Content Analysis [42] was done using data analysis software NVivo 12. For annotation (i.e., coding), we started with an annotation scheme based on the codes used in Ferraro et al. [23] (deductive). Then, we adapted those or added new (sub-)codes based on what was mentioned in the interviews (inductive). Overall, Annotator 1 iterated through all codes and annotations several times. To increase coding reliability, we had

two sessions of collective coding (Annotator 1 with Annotator 2; Annotator 1 with Annotator 3), led by Annotator 2 and Annotator 3 respectively, focusing on one interview each. These sessions lead to only minor adjustments, indicating a high level of inter-annotator agreement. The final annotation scheme includes a total of 16 general (top-level) codes.

4 RESULTS

In this section, we report our main results. First, addressing RQ1, we describe what artists deem important and fair in MRS, and what kind of improvements could be made in those areas (Subsection 4.1). We then dedicate Subsection 4.2 to what artists' considerations would be if they could control recommendations of their own music. Following this, we move on to RQ2 and highlight artists' view on music streaming platform and MRS transparency (Subsection 4.3), the role that streaming platforms currently play in artists' careers (Subsection 4.4), and the role artists envision for streaming platform in fostering fairness and diversity (Subsection 4.5). Finally, in Subsection 4.6, we outline new functionalities that artists suggested to bring their fairness improvement ideas to fruition.

We refer to participants according to their codes from Table 1. For each theme, we indicate whether there is consensus, or if artists' opinions differ. When applicable, we compare our results with those from previous studies ([23, 57]). We note here that artists indicate Spotify as the streaming platform they primarily interact with and focus on when releasing music, followed by YouTube and Apple Music. Also, we note that while most questions aim specifically at

Table 2: Topics and general scope of guiding interview questions.

No.	Topic	Question(s)
Q1	MRS experiences	What are your experiences with music streaming platforms as a user, and an artist?
Q2	Platform influence	How do music streaming platforms currently influence your career, and what was or would have been different without them?
Q3	Transparency	To what extent is it clear to you why MRS recommend which music to whom? Are you satisfied with the current transparency to artists?
Q4	Control	If you could control which of your music was recommended to whom, what would you do? Are you satisfied with the current level of control you have as an artist?
Q5	New music & repertoire	How should MRS balance newly released music and older music?
Q6	New artists	(How) should MRS promote music from artists that are just starting their careers?
Q7	Popularity (bias)	(How) should MRS promote older music that has not been played or liked often?
Q8	Localization & quotas	(How) should recommendations be localized to a user's location?
Q9	Diversity & fairness	(How) should MRS promote music from artists that are from a demographic and/or historic minority (e.g., gender)?
Q10	Influencing users	(How) should MRS try to make users' listening behavior more varied?
Q11	COVID-19 impact	Has the COVID-19 pandemic influenced your view on music streaming platforms, and the role they play in your career?
Q12	Income distribution	What is your view on the current income distribution model? Do you think songs recommended by a MRS should be worth more or less than actively selected songs?
Q13	Additions	Is there something important we did not cover yet, or you wish to re-emphasize?

experiences as an artist, sometimes artists explicitly answered from a user perspective. We indicate it when that was the case.

4.1 What Should Recommender Systems Recommend: In General

Below, we describe artists' considerations regarding MRS in general (Q1), and for different fairness dimensions explicitly: novelty (Q5 and Q6); artist popularity (Q7); localization (Q8); and gender (Q9). We also outline our observation that artists continuously keep user preferences in mind while contemplating their answers.

New artists and music. Addressing the cold start problem, we explored how streaming platforms should handle new artists and new music. Most artists encourage emphasizing acts who just started releasing music on streaming platforms. P14: "You could push quite some new bands that way, because there is simply a lot of good music being created nowadays." Still, P2, P5, and P10 note that not all artists can be famous, or even recommended at all. P10: "As a new artist... to all but count on being picked up, that does kind of take the stardom out of it. [...] It is just very difficult to become one of the greats."

Some artists want more emphasis on recommending new music, but not necessarily from new artists (P3, P4+b, P6, P7a, P14), though they note the complexity of this task. P7a: "It would be very desirable. But I'm not sure whether it is [the platforms'] task. Seeing how many new songs are added each day, that would be very tricky." P2, P3, and P5 point out not all music is worth recommending. P2: "A lot of it is simply... bad. So, being Spotify, how are you going to make a selection?"

Popularity bias. Our participants report often encountering popularity bias, as a user and as an item provider. P8a describes "What is popular will only become more popular, so that gap keeps growing.", while P5 notes "Something the algorithm doesn't know [is

recommended less], but that doesn't mean [it] is not interesting." On recommending more music from the long tail, our participants voice similar opinions as those in Ferraro et al. [23]. They generally agree that items from the long tail should be recommended more, as opposed to primarily items that are already popular.

Most artists mention recommending less popular artists as an important goal. P2: "If you could indeed add a percentage of unknown bands, or acts, [...] that would have benefits", also noting that if users want to discover new music, they might not care whether it is by a popular artist (P6), see also Ferwerda et al. [24]. However, they do not always have a clear view on how exactly that should be achieved. In addition, recommending specific older, less popular repertoire came up for P4b, P7a+b, P9, P13, and P14. P7b: "The old back catalog was imported all at once in all those databases, and it is impossible to promote it. While it might contain material that people would love to hear." Still, P1, P2, P4b, P7b, and P10 recognize that not all users are interested in items from the long tail. P10: "I know there are many people that are solely looking for really famous artists, in which case: fine, do your thing."

Location and localization. Ferraro et al. [23] contains a considerable discussion on whether or not to adapt recommendations based on users' location. They primarily discuss quotas for local music, which some artists support and others do not. In our interviews, artists in six interviews (P1, P3, P5, P6, P7a+b, P14) generally also prefer more localization, with four of them noting quotas would be a good way to achieve that, and two (P1, P14) noting they do not. P14: "[...] then, it would become something obligatory, while you don't know if you [...] have the quality and the quantity." Conversely, in four interviews (P2, P8a+b, P12, P13), participants prefer a more global approach rather than a more localized one. Lastly, P4a+b, P9, P10, and P11 are undecided, with P9 and P10 noting a quota would be good in specific cases, P9: "If it really starts to stand out that there

suddenly are a lot of Dutch artists, then it [...] can also scare people off, and it doesn't help anyone at all." Some artists also remark that, when localizing, P4a: "The music does have to be worth it."

Gender fairness. In Ferraro et al. [23], there was a clear consensus that to reach a gender balance on streaming platforms, it is warranted to explicitly promote content by artists that do not identify as men. Among our interviewees, however, there is disagreement. Some prefer not to make any selections based on gender for recommended music (P1, P3, P4a+b, P10, P12), even if that means gender balance is not reached. P1: "Look, if someone only listens to artists who are men, it probably has some reason." Their comments can generally be summarized as 'the gender of the artist should not matter when deciding who is recommended'. Some participants note that some users might prefer artists from a certain gender, and actively intervening based on this aspect could decrease user satisfaction. P12 also finds it problematic to automatically infer the singer's gender. P12: "[...] in our genre, sometimes you are not able to hear if [the singer] is a man or a woman. So... how is that being decided? And if that is [through sound analysis], it would not be correct at all. So, [...] don't do that!" Conversely, other artists feel strongly that platforms should promote content by women and non-binary artists in order to reach a balance (P2, P6, P7a+b, P8a+b, P9, P11, P14). P14: "Yes, absolutely. [...] Women create music that is just as good as music created by men." P9: "If you keep reflecting what is currently happening, the gap will only become bigger. [...] if, primarily, young boys can recognize themselves in an artist, [...] then you'll get more boys that make music, and fewer girls or other genders that think: that is something that I could do, too." P5 and P13 were undecided on the topic. P13: "Women are not some kind of helpless children, that you have to favor all the time."

One possibility to achieve gender balance, as suggested by some artists in Ferraro et al. [23], is enforcing a quota. Some of our participants are open to platforms exploring this method (P7a+b, P8a+b, P9, P11). P8a notes, "Half the world's population is made up of women: half-half. That's where it should go. Because I really don't believe that women make worse music. [...] and if you don't throw a quota on it, it won't happen. I'm convinced of that. Because, yeah, there are so many men walking around in the music world...". Still, P7b acknowledges that it might be challenging to choose and implement quotas in practice. Some artists look for alternative ways to increase gender balance as opposed to adapting algorithmic recommendations: offer playlists focusing on diversity (P6, P10); hiring curators from a diverse background (P7a+b); and encouraging women to make music in the first place (P10, P13).

Other factors. Artists consider several other factors when addressing what MRS should focus on. Firstly, as in Ferraro et al. [23], some artists (P3, P4a+b, P6, P7a+b) think that generally artists' most recent work should be recommended, as that is what represents them at that point in time. P1: "Most artists are always like: whatever I released last, is what I currently love to play the most." P1 and P10 would rather focus on the most popular songs, in order to have the greatest chance of success, though P10 notes "I would start with something that is guaranteed to be popular, and if people like it, [...] make the switch to less popular songs." The others indicate both should be possible, depending on artist preference.

For some final suggestions, P1, P6, P10, and P12 advise recommending active acts more often. Another suggestion (P5, P7, P8a+b, P10, P12) is increasing the 'human factor', e.g., through hiring more curators to improve recommendations in playlists. Lastly, P5 wants more focus on recommending artists' music to users that 'follow' or 'subscribe to' them.

Emphasis on user preference. What stands out in many interviews is the emphasis on, at minimum, studying user behavior to see whether users want more or less of a certain type of music. It is even often suggested to give users direct control to indicate, e.g., how far outside their usual listening behavior recommendations should be—a concept explored in Porcaro et al. [51], and which we further describe in Subsection 4.6. In ten interviews, artists suggest taking different user types into account, as studied in Sanna Passino et al. [54]. Interestingly, P2, P5, and P11 suggest recommending more discovery-targeted content to younger users, explaining that P2: "they are still more focused on developing their taste." Adapting MRS to user behavior is also mentioned in five interviews. P9: "If you see that users are not satisfied with [discovery-oriented recommendations], you can always decide 'OK, then we will dial it back'."

4.2 What Should Recommender Systems Recommend: For Artists' Own Music

In addition to recommendations in general, we asked artists whether they would want more control over which of their own songs are recommended to whom, and how they would use such functionality (Q4). Generally, artists would indeed value having more control, often comparing it to functionality on social media platforms such as Facebook. They experience their current lack of influence on recommendations of their music as a shortcoming of the system. Some do note P8a: "[...] how are you even supposed to know, as an artist [...]. You do have to be a marketing expert for that."

When discussing factors they would consider when using such functionality, artists distinguish between music-related and user-related attributes. Regarding the former, artists generally either prefer their newest songs (P7, P8, P10, P11, P14), their most popular songs (P3, P4a+b), or having the opportunity to choose from both (P1, P2, P12, P13). Arguments for newer songs are P1: "it simply represents better who we currently are", and audiences being able to recognize songs artists play at live shows (P14). For recommending most popular songs, artists name similar arguments as in Section 4.1 (Other factors). Artists also want to select certain playlists (P1, P4a, P10) or songs they personally prefer (P4a+b, P7a).

When considering user attributes, artists generally focus on user music preferences. P7a: "I would certainly look into specific target audiences." The main considered attributes are user age, location, taste, query or goal, and which other artists a user listens to. P2: "If you are starting to tour with a band, for example, or if you know that a band who sounds really similar is coming to the Netherlands for a show, [I would recommend my songs to users who listen to that band]."

Participants would also like to influence which other artists they are associated with on a platform, specifically referring to Spotify's 'Fans also like' section. Only P5 is satisfied with the artists shown there. Others criticize the suggestions as not being discovery-oriented enough (P1, P7b), or too focused on geographical proximity,

while artists have nothing in common otherwise (P1, P11, P13). P11: "[...] they do not even come close to the style of my music. So, only because I'm Dutch and they are Dutch." This corresponds to findings from Tofalvy and Koltai [66], who note that self-published or indie label-published artists tend to be paired with artists from their own country, regardless of genre.

Finally, P6 and P13 indicate they have no preference for their music's recommendations, P13: "[recommend] everything [to] everybody." Only P8a suggests not trying to control everything that happens to one's music once it has been released, P8a: "When you've made something, [...] you need to let go [at some point], you shouldn't want to influence it anymore."

4.3 Transparency of Platforms and Recommender Systems

In Ferraro et al. [23], artists described their desire for more transparent music streaming platforms, wanting to know especially how their integrated RS make decisions. In response to Q3, our participants largely share this sentiment (P1, P2, P3, P5, P8a+b, P10, P12), indicating that the current systems are not transparent, making it challenging for artists to respond or react to them. P1: "You're just trying random things, you have no idea why it works and why it doesn't." P2: "[Algorithmically generated playlists] sometimes feel completely random. [...] I do somewhat understand that it's a trick of the trade. [...] But it would sometimes be handier [to have more insight] because then you can respond to it." Most of them also express that they would delve into such insights if available, P10: "[...] how did people find me? [...] because they were listening to that one artist, or because they were listening to a certain flow, or..." P2, P5, P7a+b, P10, P11, and P14 do appreciate the information that is already available for them, which (in early 2022) primarily focuses on play counts and audience statistics. Only P4a+b are not sure they would use such insights, noting P4a: "but how would I benefit from that? [...] then you would really focus on creating music that users like, instead of creating music you find cool yourself." Lastly, P6 notes that even if transparency would be improved, P6: "[...] it is of no benefit to you. [...] you can't steer your audience with other behavior because Spotify does that itself. As an artist, you currently have no influence on it at all."

4.4 Current Platform Impact

In our interviews, we notice artists mostly do not view MRS as independent entities, but see such systems as being part of a bigger whole: the music streaming platforms in which they are embedded. Therefore, we need to consider how those platforms impact and influence artists currently (Q2 and Q11, described here), and whether this could be leveraged to improve fairness and diversity in MRS and the music industry as a whole (Q9 and Q10, Subsection 4.5).

Reaching an audience. First of all, artists note that streaming platforms simultaneously make it easier to share music, and more difficult for new artists to reach an audience, confirming earlier research [23]. Most artists note that these platforms make it P4a: "very straightforward to drop [all songs] to a very large network", in an instant. Conversely, P3, P6, P7a, and P8a+b acknowledge that just releasing songs on a streaming platform without investing in

additional marketing activities does not help in reaching an audience. P8b notes being somewhat disillusioned, P8b: "when we just started [adding our music to streaming platforms], I actually—yeah, maybe that was naive—expected it would have a bit more... impact." P5, P6, P8a+b, and P10 even indicate that through streaming platforms, it is more difficult for artists to connect with their audience compared to before those platforms existed. This was also noted in Siles et al. [57]. Several artists even remark that if their career would have started nowadays, they doubt whether they would have been able to break through (P5), or would have been booked for a major Dutch festival (P8a+b). Some artists also recognize that their income from selling physical records has decreased, P13: "Now, for example, people at the merch[andise] table [...] ask 'is your music available on Spotify', then you say 'yes', and then they say 'OK' and walk away again."

Playlists. The prevalence and importance of curated playlists (i.e., playlists created by an editor) is frequently mentioned. This corresponds to Siles et al. [57] but is not emphasized to the same extent in Ferraro et al. [23]. Spotify is mentioned in particular. P1, P5, P7b, and P11 note being on certain curated playlists all but guarantees thousands of plays, sometimes even per day, and the corresponding income. Therefore, they observe that artists generally focus on curated playlists more than on having their music recommended by MRS; if a track is added to a certain curated playlist, it will be algorithmically recommended later, too. P7b: "[The curator] told us: 'I can do something with [a certain song]. I will put it on a list.' And after that, that song was played a couple of million times. So it's that simple", also underlining a feeling among part of the artists that curators are effectively gatekeepers (see McKelvey and Hunt [43]). Reaching these curators can be challenging (P1, P2, P7a+b, P8a+b, P10, P12, P13). P1: "We tried to get a hold of the right people [...] but it is almost impossible to reach a person at Spotify in the first place, ever." P12 and P13 note that other parties, e.g., record labels and shareholders, are in a better position to reach and influence them than independent artists are.

Playlists are also mentioned as a cause of disconnect between users and artists. P5 and P11 note users mainly use playlists on shuffle, sometimes in the background, which P10 confirms especially for contemporary classical music. All three voice concerns that this way of consuming makes the step significantly bigger for users to visit the artist page, and eventually connect through social media elsewhere. If their music is primarily consumed through a playlist, artists risk not building their career outside of the streaming platform, P5: "Many of my colleagues have huge streaming play counts on Spotify, [...] but are not booked for any shows, have no followers on social media at all."

Narrowing diversity. Another aspect the artists consider is streaming platforms' influence on diversity—of both the created songs and the song recommendations. Regarding the former, several participants note feeling that artists' creativity is limited by the way RS work (P2, P5, P7a+b, P8a, P9). They experience that accessible songs are more often added to curated playlists, and algorithmically recommended. Therefore, they indicate feeling pressure to adopt practices to make their songs more eligible for consumption through streaming, corresponding to earlier findings [57], and note that this makes contemporary songs and albums as a whole less

diverse. Artists also experience more focus on releasing individual tracks as opposed to albums (P2, P8a, P13), and songs becoming shorter (P1, P2, P5, P9).

On users' listening behavior on streaming platforms, most artists mention that it is often passive and less curious, resulting in users' music taste generally becoming less diverse. P6: *"More and more, simply, in a bubble. [...] There is less and less tension, it makes you very passive."* Several artists describe this effect as a vicious circle P4b: *"of things that are recommended to you all the time."* P2 fears for the next generations, P3: *"[If you] only constantly push what [users] already like, what kind of generation will there be in 20, 30 years? Stiff, pigeonholing even more."* Artists generally do recognize that platforms need to cater to users' preferences to some extent so users—still creatures of habits—stay happy.

Impact of the COVID-19 pandemic. All participants feel the emphasis on streaming was amplified during the COVID-19 pandemic due to the lack of live shows, making it all the more important to focus on fairness of streaming platforms and integrated MRS. P1, P2, P6, and P13 indicate that during the pandemic, digital streaming is the only way to reach a new or existing audience, and that it has become artists' main source of income, where that used to be ticket sales for live shows. P9, P10, and P13 therefore indicate having delved deeper into utilizing streaming to reach an audience. Concerning staying relevant without being able to play live shows, P7a+b and P13 feel an increased pressure to continuously keep releasing new songs, P13: *"[...] always releasing something, show people you're still alive."* P7a: *"Indeed, it is far [...] from what you'd actually like to do. But well, it [currently] is the only platform."* Some participants, though, do not experience this pressure, either because they did not look into streaming (P3, P4a+b) or do not notice a big difference (P14).

4.5 Platform Role in Fairness and Diversity Improvement

Next, we describe insights on whether platforms should use their influence to improve fairness and diversity (Q9 and Q10). We also describe responses to Q12 regarding income distribution.

Fostering fairness. Firstly, we asked participants how they see the platforms' role in improving fairness for historically underrepresented (demographic) minorities. Generally, artists emphasize the platforms' position of power, which comes with a responsibility. P8a: *"They already have a role. [...] you can't escape that responsibility anymore, I think, as a streaming service."* They also note the positive impact streaming platforms could have if they offered more content from a diverse set of artists. P9: *"Diversity in some aspects, in turn, will contribute to diversity in other aspects. If there are artists from many age groups, you will get different music because they all grew up in other music eras."* Still, some artists note they do not expect much in this area, because P7b: *"[I don't think] the streaming service is helped by doing that. The goal is still, primarily, to make sure people do not click away, or pause their music. So everything that distracts from that [...] is counterproductive."* P13 also remarks *"Does a platform have to be completely fair anyway? Because if you would visit a music store—which everyone glorifies when discussing Spotify or something—you also have someone working there who has*

a specific taste and limited knowledge. So, [...] it is not completely correct, saying it was fairer [back] then."

Fostering diversity. Secondly, without exception, all our interviewees distinctly express that platforms should encourage users to step outside of their usual bubbles. P6: *"I think [recommending more diverse items] would be super cool to broaden some people's horizons. That, too, is... quite a challenge nowadays."* Most artists even feel diversity across genres should be fostered, to let users have more new experiences. Some artists indicate that current algorithmic recommendations lack in this respect, and platforms should take more risks, P1: *"Well, if you already listen to artist A, one can be almost 100% sure that you are also already listening to artist B, so why is [that artist] still recommended?"* Eight artists do note that the user should still be sufficiently satisfied with their recommendations because P8b: *"if the user then drops out [...] it obviously is not an improvement."* A smaller part of discovery-oriented songs might also be a solution, P12: *"[...] keep 80% of what the listener clicked on, and searched for, and 20% of 'oh, let's show this as well'."* P4a+b, P13, and P14 focus on respecting user taste even more, describing that some people prefer listening to familiar things and not discovering new music at all times, or even at all. P13: *"I don't think people are obliged to discover new music all the time, it is very nice as well to have a style or a taste, and to surrender to that."*

Income distribution. Lastly, we discuss fairness in income from streaming platforms. Most artists remark they are not satisfied with income distribution from these platforms in general. When asked whether the payout for algorithmically recommended music should be higher or lower than for music directly queried by the user, there was no clear consensus. Some artists note queried content should be worth more (P7b, P12, P14), less (P10), or all worth the same (P1, P4a+b, P11, P13), and some indicate being unsure (P6, P8a+b, P9). Still, some artists note P7a: *"If [streaming platforms] could just pay more, then a lot of issues will already be solved, I think."*

4.6 New Functionality

To finish our result outline, we note that nearly all artists suggested new functionalities for streaming platforms. According to them, those functionalities could simultaneously tackle artist fairness issues, and be welcomed by users. For instance, it could be made easier to connect artists more deeply with their audience by linking social media to their streaming platform artist profile (P5, P6, P10), though P5 notes *"it would do [streaming platforms] credit, but they will probably never do that."* Another suggestion is to offer pop-ups to users that have listened to music by an artist often but have not visited the artist's profile yet (P10). Participants also propose to make playlists designated to a certain goal, such as offering more diverse content (P3, P6, P7a+b, P9, P10, P11, P13), addressing the cold start problem (P3, P10), P3: *"Make a playlist containing only songs that have never been listened to [by anyone], but is [personalized] to some extent"*, or highlighting songs from a specific region (P7a+b, P11, P13). Another possible goal is shedding light on earlier work (P7b). One way to achieve that, suggested by P7b, is through a playlist consisting of long tail songs that were released exactly 10, 20, or 30 years ago. Finally, several artists suggest increased user control by letting users decide which songs to exclude or include in

recommendations (P7a+b) or integrating sliders or filters in the UI (P4a+b, P6, P11, P13), P11: "Just like at Zalando, [selecting yourself:] 'well, I want bands, not famous, a little bit in that genre, very new...'"

5 DISCUSSION

Here, we discuss where our results correspond to earlier findings, our additional or contradictory findings to earlier work, and we close with suggestions on how to address the various ideas for improvement.

Confirming earlier findings. On various topics, our results corroborate earlier findings [23, 57]. Our participants agree that the advent of music streaming platforms made it much easier to distribute music, though they observe that those systems make it even harder to break through as an unknown artist. *Popularity bias* is mentioned as further amplifying those difficulties for unknown artists. Platforms' influence on artists' creative process and song releasing strategy is generally recognized. Our participants also underline the importance of being added to (*curated*) playlists, which in turn increases algorithmic recommendations of their music. Regarding *transparency*, we confirm the perceived lack of MRS clarity and desire for more information about the inner workings of major platforms' MRS. Finally, the artists generally desire *more control* over how their music is recommended, and have similar considerations to those in Ferraro et al. [23].

Additional & contradictory findings. Firstly, across our interviews, artists observed that the *COVID-19 pandemic* amplified the power of platforms. They also argued that users' music taste has become less diverse, partly due to users engaging with music from specific artists less, and more through playlists. Resulting from the perceived lack of diversity, participants urged to *break users out of their bubble* and let them discover music outside their typical taste, contrary to previous research in which Spanish-speaking artists noted "[a] system should not influence the user's taste" [23]. One suggested way to increase diversity while keeping user's preferences in mind, is *giving users more active control* in which music is recommended to them. This could be through UI functionalities such as sliders and filters, or through more designated playlists with, e.g., less known artists or non-western artists. As artists feel platforms have the responsibility to use their power to foster diversity in recommendations to users and increase fairness for artists, this can be taken as advice to MRS researchers and platform developers as a whole. Still, on intervening in MRS to reach *gender balance*, the statements in our interviews were a lot more nuanced than in Ferraro et al. [23], with several artists (4 identifying as male, 1 as female) being against platforms intervening on this aspect. Participants for whom intervening based on artist gender specifically invoked resistance did not always clearly articulate the reason why they regard this aspect differently compared to other fairness dimensions. Based on their comments, we infer that these participants either view gender balance as a less important issue, or fundamentally object to using gender as a feature in RS. On localization, while Spanish-speaking artists unanimously supported promoting local content more, four Dutch artists in fact prefer a more globalized approach. Lastly, while Spanish-speaking artists noted it is important to provide music's context (e.g., the message behind songs or artists), the artists in our

sample put more emphasis on *lacking social context*; they desire more ways for users to connect with the artist behind the music, e.g., through social media or pop-ups.

While the differences between the two samples might be related to the differences regarding Hofstede's six dimensions, our approach does not allow for drawing conclusions on this aspect. Yet, in line with Knees et al. [37], we see it is crucial to consider cultural context, as adopting a study design in a different cultural context may provide new insights.

Suggestions. Based on these insights, we share five suggestions that could increase fairness, diversity and transparency:

- (1) By using active interventions in recommendations, ranging from re-ranking to implementing quotas, several fairness and diversity improvements could be realized. This would benefit artists and users alike.
- (2) By improving facilities for artists to connect with their audience, streaming platforms could contribute to a more lasting bond between artists and their listener base. This could be achieved, e.g., by connecting artists' social media accounts to their profiles, or by more actively enticing users to visit artists' profiles. Artists expect that improving such facilities would, in turn, contribute to nurturing fairness.
- (3) Offering explanations could improve MRS transparency [64], but more research is needed to understand how exactly. We expect recent advancements in governmental regulations, such as the Digital Services Act for countries in the European Union, to have a positive impact on transparency. Such frameworks oblige online platforms to "establish a powerful transparency and a clear accountability" [19].
- (4) By increasing artists' control over which of their music is recommended to whom, artists could adapt their song recommendations to their overall strategy. Novel campaign tools such as Spotify's 'Marquee'⁶, which allows a more focused music release strategy, might be the start of more progress in this respect.
- (5) Increasing users' control might encourage the exploration and discovery of new items. For example, slider functionalities ([39, 40]) or tag-based filtering might be employed to give users control over what they are recommended, and facilitate personalized exploration on dimensions beyond genre exploration.

We encourage to further explore these topics, whilst also studying the implication of such implementations.

6 CONCLUSION

In this study based on semi-structured interviews with 14 music artists, we addressed item providers' perspectives on the impact of music streaming platforms, and explored their considerations on MRS fairness. Our contribution is three-fold. First, for a set of fairness aspects, our results confirm earlier findings [23], yet in another cultural context. Second, our study contributes valuable additional insights to the previous study, with more nuanced responses on some fairness aspects (e.g., gender balance, localization). It is important to more closely explore those aspects. Third, our work

⁶<https://artists.spotify.com/marquee>

contributes concrete ideas for fairness and user control-enhancing UI functionalities on music streaming platforms.

Our study comes with a set of limitations, the first one relating to focusing on artists from one country. While a limitation in terms of generalization, it makes a clear comparison with different cultures in past and future work possible. Further, this study relied upon a volunteer sample, raising questions about who chose to volunteer and who did not [53]. Moreover, the interview setting could have resulted in social desirability response bias: a tendency for participants to frame their views in ways they believed were expected from or valued by the interviewer [41, 52].

In terms of future work, our suggestions on both UI and how fairness aspects should be implemented in recommendation algorithms need further research. Thereby, it is essential to not only address those elements' design, but also to evaluate whether these fulfill the needs as expected. This will also require longitudinal studies to assess long-term impact. Inspired by Thorndike [63], the ultimate goal should not (only) be to have fairer systems, but that these impact the real world: for instance, in terms of how often artists' songs are actually played.

ACKNOWLEDGMENTS

We thank all participants for sharing their invaluable insights. We also express our gratitude to all who supported us in recruiting participants. Lastly, we thank Judith Masthoff for her insightful feedback and suggestions.

REFERENCES

- [1] Age Discrimination in Employment Act. 1967. Age Discrimination in Employment Act of 1967.
- [2] Christine Bauer, Marta Kholodylo, and Christine Strauss. 2017. Music Recommender Systems: Challenges and Opportunities for Non-Superstar Artists. In *Proceedings of 30th Bled eConference* (Bled, Slovenia), Andreja Pucihar, Mirjana Kljajić Borštnar, Christian Kittl, Pascal Ravesteijn, Roger Clarke, and Roger Bons (Eds.). University of Maribor Press, Maribor, Slovenia, 21–32. <https://doi.org/10.18690/978-961-286-043-1>
- [3] Ludovico Boratto, Gianni Fenu, and Mirko Marras. 2021. Interplay between upsampling and regularization for provider fairness in recommender systems. *User Modeling and User-Adapted Interaction* 31, 3 (2021), 421–455. <https://doi.org/10.1007/s11257-021-09294-8>
- [4] Ludovico Boratto, Gianni Fenu, Mirko Marras, and Giacomo Medda. 2022. Consumer Fairness in Recommender Systems: Contextualizing Definitions and Mitigations. In *Proceedings of the European Conference on Information Retrieval* (Stavanger, Norway) (*ECIR '22*), Matthias Hagen, Suzan Verberne, Craig Macdonald, Christin Seifert, Krisztian Balog, Kjetil Nørkvåg, and Vinay Setty (Eds.). Springer International Publishing, Cham, Germany, 552–566. https://doi.org/10.1007/978-3-030-99736-6_37
- [5] Georgina Born, Jeremy Morris, Fernando Diaz, and Ashton Anderson. 2021. *Artificial Intelligence, Music Recommendation, and the Curation of Culture*. White Paper. A. Schwartz Reisman Institute for Technology and Society.
- [6] Theodor Cimpanu, Alessandro Di Stefano, Cedric Perret, and The Anh Han. 2023. Social diversity reduces the complexity and cost of fostering fairness. *Chaos, Solitons & Fractals* 167, Article 113051 (2023), 9 pages. <https://doi.org/10.1016/j.chaos.2022.113051>
- [7] Civil Rights Act. 1964. Civil Rights Act of 1964.
- [8] Patricia Hill Collins and Sirma Bilge. 2020. *Intersectionality*. John Wiley & Sons, Chichester, UK.
- [9] John W Creswell and Cheryl N Poth. 2016. *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications, Thousand Oaks, CA, USA.
- [10] Fernando Diaz, Bhaskar Mitra, Michael D. Ekstrand, Asia J. Biega, and Ben Carterette. 2020. Evaluating Stochastic Rankings with Expected Exposure. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (Virtual Event, Ireland) (*CIKM '20*). ACM, New York, NY, USA, 275–284. <https://doi.org/10.1145/3340531.3411962>
- [11] Karlijn Dinnissen and Christine Bauer. 2022. Fairness in music recommender systems: a stakeholder-centered mini review. *Frontiers in Big Data* 5, Article 913608 (2022), 9 pages. <https://doi.org/10.3389/fdata.2022.913608>
- [12] Vasco Eiriz and Filipe Pinto Leite. 2017. The digital distribution of music and its impact on the business models of independent musicians. *The Service Industries Journal* 37, 13–14 (2017), 875–895. <https://doi.org/10.1080/02642069.2017.1361935> arXiv:<https://doi.org/10.1080/02642069.2017.1361935>
- [13] Michael D Ekstrand, Anubrata Das, Robin Burke, and Fernando Diaz. 2022. Fairness in Information Access Systems. *Foundations and Trends® in Information Retrieval* 16, 1–2 (2022), 177 pages. <https://doi.org/10.1561/15000000079>
- [14] Michael D. Ekstrand, Mucun Tian, Ion Madrazo Azpiazu, Jennifer D. Ekstrand, Oghenemaro Anuyah, David McNeill, and Maria Soledad Pera. 2018. All The Cool Kids, How Do They Fit In?: Popularity and Demographic Biases in Recommender Evaluation and Effectiveness. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* (Nice, France) (*FAT* '18*, Vol. 81), Sorelle A. Friedler and Christo Wilson (Eds.). PMLR, New York, NY, USA, 172–186. <https://proceedings.mlr.press/v81/ekstrand18b.html> <http://proceedings.mlr.press/v81/ekstrand18b/ekstrand18b.pdf>
- [15] Michael D. Ekstrand, Mucun Tian, Mohammed R. Imran Kazi, Hoda Mehrpouyan, and Daniel Kluver. 2018. Exploring Author Gender in Book Rating and Recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems* (Vancouver, British Columbia, Canada) (*RecSys '18*). ACM, New York, NY, USA, 242–250. <https://doi.org/10.1145/3240323.3240373>
- [16] Mehdi Elahi, Danial Khosh Kholgh, Mohammad Sina Kiarostami, Soroush Saghari, Shiva Parsa Rad, and Marko Tkalcic. 2021. Investigating the impact of recommender systems on user-based and item-based popularity bias. *Information Processing & Management* 58, 5, Article 102655 (2021), 15 pages. <https://doi.org/10.1016/j.ipm.2021.102655>
- [17] Avriel Epps-Darling, Romain Takeo Bouyer, and Henriette Cramer. 2020. Artist Gender Representation in Music Streaming. In *Proceedings of the 21st International Society for Music Information Retrieval Conference* (Montreal, Canada, October 11–16) (*ISMIR '20*). 248–254. <https://archives.ismir.net/ismir2020/paper/000148.pdf>
- [18] European Union. 2010. *Charter of Fundamental Rights of the European Union*. Vol. 53. European Union, Brussels. 380 pages.
- [19] European Union. 2022. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act). <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>
- [20] Andres Ferraro, Gustavo Ferreira, Fernando Diaz, and Georgina Born. 2022. Measuring Commonality in Recommendation of Cultural Content: Recommender Systems to Enhance Cultural Citizenship. In *Proceedings of the 16th ACM Conference on Recommender Systems* (Seattle, WA, USA) (*RecSys '22*). ACM, New York, NY, USA, 567–572. <https://doi.org/10.1145/3523227.3551476>
- [21] Andrés Ferraro, Jea Ho Jeon, Biho Kim, Xavier Serra, and Dmitry Bogdanov. 2020. Artist biases in collaborative filtering for music recommendation. In *Proceedings of the 37th International Conference on Machine Learning* (Vienna, Austria) (*ICML '20*). ICML, PMLR, New York, NY, USA, 1–3. <http://hdl.handle.net/10230/45185>
- [22] Andrés Ferraro, Xavier Serra, and Christine Bauer. 2021. Break the Loop: Gender Imbalance in Music Recommenders. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval* (Virtual Event) (*CHIIR '21*). ACM, New York, NY, USA, 249–254. <https://doi.org/10.1145/3406522.3446033>
- [23] Andrés Ferraro, Xavier Serra, and Christine Bauer. 2021. What Is Fair? Exploring the Artists' Perspective on the Fairness of Music Streaming Platforms. In *Human-Computer Interaction – INTERACT 2021: 18th IFIP TC 13 International Conference* (Bari, Italy) (*INTERACT '21*, Vol. 12933), C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, and K. Inkpen (Eds.). Springer, Cham, Germany, 562–584. https://doi.org/10.1007/978-3-030-85616-8_33
- [24] Bruce Ferwerda, Michaela Berndt, Eveline Ingesson, and Markus Schedl. 2023. I Don't Care How Popular You Are! Investigating Popularity Bias in Music Recommendations from a User's Perspective. In *ACM SIGIR Conference on Human Information Interaction and Retrieval* (Austin, TX, USA) (*CHIIR '23*). ACM, New York, NY, USA, 357–361. <https://doi.org/10.1145/3576840.3578287>
- [25] Greg Guest, Arwen Bunce, and Laura Johnson. 2006. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods* 18, 1 (2006), 59–82. <https://doi.org/10.1177/1525822X05279903>
- [26] Jaron Harambam, Dimitrios Bountouridis, Mykola Makhortykh, and Joris van Hoboken. 2019. Designing for the Better by Taking Users into Account: A Qualitative Evaluation of User Control Mechanisms in (News) Recommender Systems. In *Proceedings of the 13th ACM Conference on Recommender Systems* (Copenhagen, Denmark) (*RecSys '19*). ACM, New York, NY, USA, 69–77. <https://doi.org/10.1145/3298689.3347014>
- [27] Natali Helberger, Theo Araujo, and Claes H. de Vreese. 2020. Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making. *Computer Law & Security Review* 39, Article 105456 (2020), 16 pages. <https://doi.org/10.1016/j.clsr.2020.105456>
- [28] Karla Hernandez, Stacy L. Smith, Marc Choueiti, and Katherine Pieper. 2022. *Inclusion in the Recording Studio?: Gender and Race/Ethnicity of Artists, Songwriters*

- & Producers across 1,000 Popular Songs from 2012–2021. Technical Report. Annenberg Inclusion Initiative. <https://assets.uscannenberg.org/docs/aii-inclusion-recording-studio-20220331.pdf>.
- [29] Geert Hofstede, Gert Jan Hofstede, and Michael Minkov. 2010. *Cultures and organizations: Software of the mind*. Vol. 3rd, revised. McGraw-Hill, New York, NY, USA.
 - [30] Geert Hofstede et al. 2015. Dimension Data Matrix [data]. <https://geerthofstede.com/research-and-vsm/dimension-data-matrix/>
 - [31] Ben Hutchinson and Margaret Mitchell. 2019. 50 Years of Test (Un)Fairness: Lessons for Machine Learning. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) (FAT* '19). ACM, New York, NY, USA, 49–58. <https://doi.org/10.1145/3287560.3287600>
 - [32] Dietmar Jannach and Christine Bauer. 2020. Escaping the McNamara Fallacy: Towards more Impactful Recommender Systems Research. *AI Magazine* 41, 4 (Dec 2020), 79–95. <https://doi.org/10.1609/aimag.v41i4.5312>
 - [33] Dietmar Jannach, Paul Resnick, Alexander Tuzhilin, and Markus Zanker. 2016. Recommender Systems—Beyond Matrix Completion. *Commun. ACM* 59, 11 (oct 2016), 94–102. <https://doi.org/10.1145/2891406>
 - [34] Olena Khlystova, Yelena Kalyuzhnova, and Maksim Belitski. 2022. The impact of the COVID-19 pandemic on the creative industries: A literature review and future research agenda. *Journal of Business Research* 139 (feb 2022), 1192–1210. <https://doi.org/10.1016/j.jbusres.2021.09.062>
 - [35] Sungchan Kim and Soyoung Park. 2017. Diversity Management and Fairness in Public Organizations. *Public Organization Review* 17, 2 (2017), 179–193. <https://doi.org/10.1007/s11115-015-0334-y>
 - [36] Peter Knees, Andrés Ferraro, and Mortitz Hübler. 2022. Bias and Feedback Loops in Music Recommendation: Studies on Record Label Impact. In *Proceedings of the 2nd Workshop on Multi-Objective Recommender Systems, co-located with 16th ACM Conference on Recommender Systems (RecSys 2022)* (Seattle, WA, USA) (CEUR Workshop Proceedings, Vol. 3268). CEUR-WS.org, Aachen, Germany, Article 6, 10 pages. <https://ceur-ws.org/Vol-3268/paper6.pdf>
 - [37] Peter Knees, Bruce Ferwerda, Andreas Rauber, Sebastian Strumbelj, Annabel Resch, Laurenz Tomandl, Valentin Bauer, Tang Fung Yee, Josip Bobinac, Amila Ceranic, and Riad Dizdar. 2022. A Reproducibility Study on User-centric MIR Research and Why it is Important. In *Proceedings of the 23rd International Society for Music Information Retrieval Conference 2022* (Bengaluru, India) (ISMIR 2022), 8 pages.
 - [38] Dominik Kowald, Peter Müllner, Eva Zangerle, Christine Bauer, Markus Schedl, and Elisabeth Lex. 2021. Support the underground: Characteristics of beyond-mainstream music listeners. *EPJ Data Science* 10, 1, Article 14 (2021), 28 pages. <https://doi.org/10.1140/epjds/s13688-021-00268-9>
 - [39] Yu Liang and Martijn C. Willemsen. 2021. The Role of Preference Consistency, Defaults and Musical Expertise in Users' Exploration Behavior in a Genre Exploration Recommender. In *Proceedings of the 15th ACM Conference on Recommender Systems* (Amsterdam, The Netherlands) (RecSys '21). ACM, New York, NY, USA, 230–240. <https://doi.org/10.1145/3460231.3474253>
 - [40] Yu Liang and Martijn C. Willemsen. 2022. Exploring the Longitudinal Effects of Nudging on Users' Music Genre Exploration Behavior and Listening Preferences. In *Proceedings of the 16th ACM Conference on Recommender Systems* (Seattle, WA, USA) (RecSys '22). ACM, New York, NY, USA, 3–13. <https://doi.org/10.1145/3523227.3546772>
 - [41] Mingnan Liu, Frederick G. Conrad, and Sunghee Lee. 2017. Comparing acquiescent and extreme response styles in face-to-face and web surveys. *Quality & Quantity* 51, 2 (2017), 941–958. <https://doi.org/10.1007/s11135-016-0320-7>
 - [42] Philipp Mayring. 2004. *Qualitative content analysis* (2 ed.). Vol. 1. Sage, Thousand Oaks, CA, USA, 159–176.
 - [43] Fenwick McKelvey and Robert Hunt. 2019. Discoverability: Toward a Definition of Content Discovery Through Platforms. *Social Media + Society* 5, 1 (2019), 15 pages. <https://doi.org/10.1177/2056305118819188>
 - [44] Rishabh Mehrotra, Niannan Xue, and Mounia Lalmas. 2020. Bandit Based Optimization of Multiple Objectives on a Music Streaming Platform. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Virtual Event, CA, USA) (KDD '20). ACM, New York, NY, USA, 3224–3233. <https://doi.org/10.1145/3394486.3403374>
 - [45] Silvia Milano, Mariarosaria Taddeo, and Luciano Floridi. 2020. Recommender systems and their ethical challenges. *AI & Society* 35, 4 (2020), 957–967. <https://doi.org/10.1007/s00146-020-00950-y>
 - [46] Margaret Mitchell, Dylan Baker, Nyalleng Moorosi, Emily Denton, Ben Hutchinson, Alex Hanna, Timnit Gebru, and Jamie Morgenstern. 2020. Diversity and Inclusion Metrics in Subset Selection. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (New York, NY, USA) (AI/ES '20). ACM, New York, NY, USA, 117–123. <https://doi.org/10.1145/3375627.3375832>
 - [47] Janice M. Morse. 1994. Designing funded qualitative research. In *Handbook of qualitative research*. Sage Publications, Thousand Oaks, CA, USA, 220–235.
 - [48] Marcus O'Dair and Andrew Fry. 2020. Beyond the black box in music streaming: the impact of recommendation systems upon artists. *Popular Communication* 18, 1 (2020), 65–77. <https://doi.org/10.1080/15405702.2019.1627548>
 - [49] Ricardo S Oliveira, Caio Nóbrega, Leandro Balby Marinho, and Nazareno Andrade. 2017. A Multiobjective Music Recommendation Approach for Aspect-Based Diversification. In *Proceedings of the 18th International Society for Music Information Retrieval Conference* (Suzhou, China) (ISMIR '17), 414–420. <https://archives.ismir.net/ismir2017/paper/000153.pdf>
 - [50] Evaggelia Pitoura. 2020. Social-Minded Measures of Data Quality: Fairness, Diversity, and Lack of Bias. *ACM Journal of Data and Information Quality* 12, 3, Article 12 (jul 2020), 8 pages. <https://doi.org/10.1145/3404193>
 - [51] Lorenzo Porcaro, Emilia Gómez, and Carlos Castillo. 2022. Diversity in the Music Listening Experience: Insights from Focus Group Interviews. In *ACM SIGIR Conference on Human Information Interaction and Retrieval* (Regensburg, Germany) (CHIIR '22). ACM, New York, NY, USA, 272–276. <https://doi.org/10.1145/3498366.3505778>
 - [52] Donna M Randall and Maria F Fernandes. 1991. The social desirability response bias in ethics research. *Journal of business ethics* 10, 11 (1991), 805–817.
 - [53] Robert Rosenthal. 1965. The volunteer subject. *Human Relations* 18, 4 (1965), 389–406.
 - [54] Francesco Sanna Passino, Lucas Maystre, Dmitrii Moor, Ashton Anderson, and Mounia Lalmas. 2021. Where To Next? A Dynamic Model of User Preferences. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) (WWW '21). ACM, New York, NY, USA, 3210–3220. <https://doi.org/10.1145/3442381.3450028>
 - [55] Markus Schedl, Peter Knees, Brian McFee, and Dmitry Bogdanov. 2022. Music Recommendation Systems: Techniques, Use Cases, and Challenges. In *Recommender Systems Handbook* (3rd ed.), Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer US, New York, NY, USA, 927–971. https://doi.org/10.1007/978-1-0716-2197-4_24
 - [56] Andrew I. Schein, Alexandrin Popescu, Lyle H. Ungar, and David M. Pennock. 2002. Methods and Metrics for Cold-Start Recommendations. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (Tampere, Finland) (SIGIR '02). ACM, New York, NY, USA, 253–260. <https://doi.org/10.1145/564376.564421>
 - [57] Ignacio Siles, Amy Ross Arguedas, Mónica Sancho, and Ricardo Solís-Quesada. 2022. Playing Spotify's game: artists' approaches to playlisting in Latin America. *Journal of Cultural Economy* 15, 5 (2022), 17 pages. <https://doi.org/10.1080/17530350.2022.2058061>
 - [58] Jaewung Sim, Daegon Cho, Youngdeok Hwang, and Rahul Telang. 2022. Frontiers: Virus Shook the Streaming Star: Estimating the COVID-19 Impact on Music Consumption. *Marketing Science* 41, 1 (2022), 19–32. <https://doi.org/10.1287/mksc.2021.1321> arXiv:https://doi.org/10.1287/mksc.2021.1321
 - [59] Nasim Sonboli, Robin Burke, Michael Ekstrand, and Rishabh Mehrotra. 2022. The Multisided Complexity of Fairness in Recommender Systems. *AI Magazine* 43, 2 (Jun. 2022), 164–176. <https://doi.org/10.1002/aaai.12054>
 - [60] Nasim Sonboli, Jessie J. Smith, Florencia Cabral Berenfus, Robin Burke, and Casey Fiesler. 2021. Fairness and Transparency in Recommendation: The Users' Perspective. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* (Utrecht, The Netherlands) (UMAP '21). ACM, New York, NY, USA, 274–279. <https://doi.org/10.1145/3450613.3456835>
 - [61] Sasha Stoikov and Hongyi Wen. 2021. Evaluating Music Recommendations with Binary Feedback for Multiple Stakeholders. In *Proceedings of the 1st Workshop on Multi-Objective Recommender Systems (MORS 2021) co-located with 15th ACM Conference on Recommender Systems (RecSys 2021)* (Amsterdam, The Netherlands) (MORS '21). CEUR Workshop Proceedings, ACM, New York, NY, USA, Article 9, 7 pages. <http://ceur-ws.org/Vol-2959/paper9.pdf>
 - [62] Tom Sühr, Sophie Hilgard, and Himabindu Lakkaraju. 2021. Does Fair Ranking Improve Minority Outcomes? Understanding the Interplay of Human and Algorithmic Biases in Online Hiring. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) (AI/ES '21). ACM, New York, NY, USA, 989–999. <https://doi.org/10.1145/3461702.3462602>
 - [63] Robert L Thorndike. 1971. Concepts of culture-fairness. *Journal of Educational Measurement* 8, 2 (1971), 63–70.
 - [64] Nava Tintarev and Judith Masthoff. 2022. Beyond Explaining Single Item Recommendations. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer US, New York, NY, 711–756. https://doi.org/10.1007/978-1-0716-2197-4_19
 - [65] Nava Tintarev, Emily Sullivan, Dror Guldin, Sihang Qiu, and Daan Odijk. 2018. Same, Same, but Different: Algorithmic Diversification of Viewpoints in News. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization* (Singapore) (UMAP '18). ACM, New York, NY, USA, 7–13. <https://doi.org/10.1145/3213586.3226203>
 - [66] Tamas Tofalvy and Júlia Koltai. 2021. "Splendid Isolation": The reproduction of music industry inequalities in Spotify's recommendation system. *New Media & Society* (2021), 1461448211022161. <https://doi.org/10.1177/1461448211022161>
 - [67] Peter Tschmuck. 2021. *The Economics of Music* (2nd ed.). Agenda Publishing, Newcastle upon Tyne, UK. <https://doi.org/10.2307/j.ctv1wgvb9x>
 - [68] Yixue Wang and Emőke-Ágnes Horvát. 2019. Gender Differences in the Global Music Industry: Evidence from MusicBrainz and The Echo Nest. In *Proceedings*

- of the 13th International AAAI Conference on Web and Social Media (Munich, Germany) (ICWSM '19, 01). Association for the Advancement of Artificial Intelligence, Palo Alto, CA, 517–526. <https://ojs.aaai.org/index.php/ICWSM/article/view/3249>.
- [69] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. 2022. A Survey on the Fairness of Recommender Systems. *ACM Transactions on Information Systems* 41, 3, Article 52 (jun 2022), 43 pages. <https://doi.org/10.1145/3547333>
- [70] Jada Watson. 2020. Programming Inequality: Gender Representation on Canadian Country Radio (2005–2019). In *Proceedings of the 21st International Society for Music Information Retrieval Conference* (Montréal, Canada) (ISMIR 2020), 392–399.
- [71] Eva Zangerle and Christine Bauer. 2022. Evaluating recommender systems: survey and framework. *Comput. Surveys* 55, 8, Article 170 (2022), 38 pages. <https://doi.org/10.1145/3556536>