Porcaro, L., et al. (2021). Diversity by Design in Music Recommender Systems. *Transactions of the International Society for Music Information Retrieval,* 4(1), pp. 114–126. DOI: https://doi.org/10.5334/tismir.106

OVERVIEW ARTICLE

Diversity by Design in Music Recommender Systems

Lorenzo Porcaro*, Carlos Castillo† and Emilia Gómez*‡

Music Recommender Systems (Music RS) are nowadays pivotal in shaping the listening experience of people all around the world. Partly driven by the commercial application of this technology, music recommendation research has gained increasing attention both within and outside the Music Information Retrieval (MIR) community. Thanks also to the widespread use of recommender systems in music streaming services, it has been possible to enhance several characteristics of such systems in terms of performance, design, and user experience. Nonetheless, imagining Music RS only from an application-driven perspective may generate an incomplete view of how this technology is affecting people's habitus, from the decision-making processes to the formation of musical taste and opinions. In this overview, we address the concept of *diversity* in music recommendation, and taking a value-driven approach we review diversity-related methodologies proposed in the Music RS literature. Additionally, by taking as an example the wider context of Information Technology (IT), we present the elements interacting in the diversity by design paradigm. We do that to acknowledge the lack of a comprehensive framework in Music RS research to address diversity, until now mostly driven by empirical results and fragmented in different application areas. Maintaining an interdisciplinary perspective, we discuss some challenges that MIR practitioners may face when researching Music RS, going beyond the search for better performance and instead questioning the theoretical foundations on which to base future research.

Keywords: Music Recommender Systems; Diversity; Information Technology

1. Introduction

Music, if conceived as a common heritage of humanity, is a heterogeneous mixture of creative processes taking shape in different historical, cultural, and societal contexts. In everyday life, cultural differences are experienced when interactions among cultures appear, leading the individual to elaborate the notion of *self* (what is familiar) in contrast to the *other* (what is different) (Wagner and Veloso, 2019). As discussed by Grenier (1989), throughout history the concept of diversity has evolved and its evolution has been fundamental in shaping relationships between different musical traditions. Whilst fields such as Musicology, Music Cognition, or Psychology and Sociology of Music have a long tradition of questioning the significance of *cultural diversity* (Huron, 2004), the younger field of Music Information Retrieval is still in its early steps in addressing similar questions, and in translating such knowledge through the design of MIR technologies accountable for social values, like *diversity* (Serra et al., 2013).

Diversifying MIR is a goal to be accomplished by understanding the multidimensionality inherent in both music and human nature. Aspects such as the diversity of the teams engaged in the design and development of MIR systems, the diversity of musical works and their creators, how to diversify tools to help MIR practitioners address cultural differences, and who and how is benefiting from the diversification strategies, and who is not, are part of the challenges described by Born (2020). Those challenges are similarly identified in the broader field of Artificial Intelligence (AI), the parent of concepts such as 'Music Intelligence' and 'AI Music' (Liebman and Stone, 2020), a field in which we are already witnessing a diversity crisis, for instance with regards to the workforce involved in the design of AI systems (West et al., 2019), or the academic community participating in AI conferences (Freire et al., 2021).

In this overview, we explore the literature related to diversity in Music Recommender Systems. Rather than focusing on the comparison of the works trying to identify 'good' or 'bad' practices, we choose to connect them arguing about *if* and *how* diversity has been embedded in Music RS, and what could be the consequences of specific designs and implementations.[1] In order to do this, in **Section 2** we take as reference the field of Information Technology, to which RS belong. We examine how diversity can be included in the design process of such technology in a principled and comprehensive manner (Friedman et

* Music Technology Group, Universitat Pompeu Fabra, Barcelona, ES

† Web Science and Social Computing Research Group, Universitat Pompeu Fabra, Barcelona, ES

‡ Joint Research Centre, European Commission, Seville, ES

Corresponding author: Lorenzo Porcaro (lorenzo.porcaro@upf.edu)

al., 2013), creating architectures which may help people in making diverse choices (Helberger, 2011), the idea defining the *diversity by design* approach. Until now, RS research has mostly focused on the improvement of performance by developing more and more sophisticated techniques, but the impact of these technologies is still underexplored (Jannach and Bauer, 2020). As highlighted by Salamon (2019), music recommendation research can be considered among the few exceptions in MIR in which there is an effective connection between research and end-users. This leaves room for questioning how such research and resulting technology are benefiting, or not, those who are actively engaged in its consumption. In **Section 3**, we review the diversity-related Music RS literature, the core part of this overview. In today's digital spaces, listening experiences cannot be imagined without considering the widespread use of streaming services, in which Music RS play a crucial role in helping people find what they want to listen to, but also in driving them by proposing music when they do not know what to choose (Schedl et al., 2018). The mere fact that streaming services in 2019 generated almost half of the global revenues in the music industry should be enough to understand the potential impact of Music RS, at least from a commercial perspective.[2] In **Section 4** we identify future challenges for the design of diversity-aware Music RS. In particular, we present findings from different disciplines which can provide to MIR practitioners new perspectives for integrating diversity not exclusively from a computational viewpoint. Lastly, we draw conclusions in **Section 5**.

At the time of writing, surveys on music recommendation diversity have not yet been presented, but in the following we present related surveys that may help the reader in deepening aspects not fully covered here. Castells et al. (2015) review evaluation procedures, algorithmic solutions, and empirical results connected with the notions of diversity and novelty in RS research. Diversity-related metrics are also discussed by Kaminskas and Bridge (2016), together with other beyond-accuracy objectives proposed in the literature. Kunaver and Požrl (2017) present an overview of RS diversification techniques, focusing both on algorithmic solutions and evaluation practices. From a wider perspective, Drosou et al. (2017) discuss the role of diversity in Big Data applications, focusing on the selection task. These surveys comprehensively treat most of the algorithmic approaches proposed in the diversity-related literature, applicable as "off-the-shelf" methods also in MIR, and therefore we encourage their reading for practitioners interested in exploring technological solutions.

Instead, the goal of this overview is threefold: 1) to present the diversity by design paradigm in IT, and its implications for Music RS research; 2) to review the MIR literature discussing proposed approaches considering diversity in Music RS; 3) to identify open challenges for the design of diversity-aware music recommendations.

2. Diversity by Design in IT

Increasing attention to value sensitive design (Friedman et al., 2013) in IT, and similarly in AI, emerged with the widespread introduction of such technologies in our daily life, and search engines and recommender systems are tangible proof (Baeza-Yates, 2018). Questions about a spectrum of topics, such as ethics, autonomy, fairness, or bias, revived the debate around practices of embedding human values and attributes in machines (Friedman and Nissenbaum, 1996). However, its relevance in the contemporary social system is undoubtedly increased as a consequence of the prominence and ubiquity of those technologies, nowadays central parts of our lives (Barocas and Selbst, 2014; Gómez et al., 2021). Throughout this overview, we consider diversity as the core principle that we want to incorporate in the value-sensitive design of Music RS. Indeed, preserving and supporting the multitude of musical languages and artistic expressions created and experienced by people all around the world is one of the goals that IT should pursue in the music domain (Serra, 2011).

The conceptualisation and measurement of diversity have been the object of study of a broad range of disciplines (Stirling, 2007; Steel et al., 2018; Mitchell et al., 2020), and to identify a global framework for measuring music recommendation diversity is out of the scope of this overview. Several diversity indexes have been formulated to describe different kinds of populations, among which most of them fall within the category of the so-called *dual-concept* diversity (McDonald and Dimmick, 2003). Such indexes make use of two dimensions: *variety*, the number of categories in a population, and *balance*, representing the evenness of elements' distribution across categories. For example, in a set of tracks classified with regards to their music genre, the variety is the number of genres within the set, while the balance is how tracks are distributed among the genres. Among the others, the Shannon, Simpson, and Herfindahl indexes fall within this category, widely used also in the MIR literature.

One of the main drawbacks of applying the dual-concept logic in the music domain is its frequentist definition of diversity, where conclusions are built only by observing the distribution of the data. Indeed, by exclusively using the variety and balance between elements and categories of a set, additional information about the nature of categories is discarded. Again in the case of tracks classified by genres, having a set of tracks half *Blues* and half *Rock*, and another one split equally into *Blues* and *Electronic* tracks, using a dual-concept diversity index the two sets may appear equally diverse in terms of track genre, while musically speaking in the former set we imagine that tracks could be less diverse, the two genres being closer than in the latter case. To overcome this issue, further dimensions describing differences between categories can be considered, as in the case of *disparity* for the Rao-Stirling index (Stirling, 2007).

From a wider perspective, the diversity by design paradigm is not just a matter of choosing the right metric, instead it is '[…] the idea that it is possible to create an architecture or service that helps people to make diverse choices' (Helberger, 2011). Diversity interplays in different aspects of the design process of information systems, and in the next section we deepen the role of those in this process. To facilitate the reading, in **Table 1** we summarise the concepts discussed in the overview.

2.1 Deconstruction, purpose and impact

Approaching diversity from an information perspective, a first step is to understand how to deconstruct this concept. Napoli (1999) identifies three components cooperating in the design of IT: 1) *source diversity*, aspects related to the information providers; 2) *content diversity*, describing the composition of the information accessible to users; 3) *exposure diversity*, identifying what content users access in contrast to what is available. In the case of Music RS, content diversity can refer to the catalogue from which recommendations are provided, source diversity to the artists or record labels providing such catalogue, and exposure diversity relates to the recommendations that listeners eventually consume.[3]

Secondly, we can identify the purpose of introducing the diversity by design approach. In broader terms, the goal is building systems that can guarantee people to be aware of the range of accessible information (Helberger, 2011). However, the motivation behind this choice may be not unique. Helberger et al. (2018) identify three perspectives: individual autonomy, deliberative and adversarial. Under an *individual autonomy perspective*, the idea is to give individuals a tool to exploit their different interests. In this case, we imagine Music RS helping people in diversifying the listening experience, broadening the possible choices with regards to their music preferences. Pursuing a *deliberative perspective*, the aim is to promote the public debate, showing divergent opinions and helping people in constructing a critical view. Here, Music RS can be designed to make listeners explore music far from their preferences, to make them aware of the unknown parts of the musical panorama. With an *adversarial perspective*, the focus is to broaden the debate highlighting non-dominant visions. Similar to the previous case, Music RS can serve as a way to promote underrepresented groups, whether subcultures or non-mainstream musical styles, under a non-hegemonic view.

Finally, we want to understand the consequences of the presence of diversity, or the lack thereof. Positive benefits of implementing diversity policies can be several, starting from fostering innovation and creativity in the workplace (Stirling, 2007), to promote equality in access to knowledge and freedom of expression (UNESCO, 2001). Such benefits hold true in the MIR field. Where diversity is lacking, damaging effects having a negative impact on people and society have already been found in IT. Among those, the phenomenon of being continuously over-exposed to content that fits our interests, named the *filter bubble* by Pariser (2011), is probably the most researched and discussed both within and outside the academic community. Similarly, *echo chambers* have been identified where technologies exacerbate the tendency to relate mainly with people with like-minded opinions (Sunstein, 2001). From a societal view, *balkanisation* refers to the fragmentation of digital spaces into different communities based on their interests (Van Alstyne and Brynjolfsson, 2005).

Under this lens, it is possible to identify the role that Music RS play in determining the exposure to music, and how most of the research until now has focused on empowering exposure diversity under an individual autonomy perspective (Helberger et al., 2018). This may be linked with the emergence of filter bubbles and echo chambers created by Music RS, wherein adversarial or deliberative perspectives could help in alleviating such

Table 1: Summary of terms and definitions presented in the overview.

Terms & Definitions	Reference(s)
Cultural diversity: the uniqueness and plurality of the identities of the groups and societies making up humankind.	UNESCO (2001) Huron (2004)
Dual-concept diversity: measurement of diversity based on the variety and balance of the elements of a population divided into categories. Variety: number of categories in a population. Balance: distribution of elements over the categories of a population. Disparity: differences between categories of a population.	McDonald and Dimmick (2003) Stirling (2007)
Diversity by design: the creation of an architecture or service that helps people to make diverse choices. Source diversity: the range of information providers. Content diversity: the range of information provided. Exposure diversity: the range of information accessed by people. Individual autonomy perspective: provide people with a tool for exploiting their different interests. Deliberative perspective: promote public awareness by showing divergent opinions. Adversarial perspective: enhance the visibility of underrepresented opinions.	Napoli (1999) Helberger (2011) Helberger el al. (2018) Loecherbach et al. (2020)
Diversity-aware RS: recommender systems designed to diversify the users' experience. Item diversity: the range of items recommended by a RS. User diversity: the range of users interacting with a RS. (User) behavioural diversity: the range of items accessed by the users. (User) perceived diversity: the item diversity as perceived by the users.	Castells et al. (2015) Kaminskas and Bridge (2016) Kunaver and Požrl (2017)

negative impact valorising underground artists, as recently explored by Kowald et al. (2021).

Whilst the areas in which such phenomena have been studied range from political views (Bozdag and van den Hovan, 2015), access to news (Lunardi, 2019), social data (Olteanu et al., 2016), and cultural products (Nguyen et al., 2014), in the field of MIR they are still underexplored. Naturally, ethical considerations on the misuses of Music RS have already emerged in the MIR community (Holzapfel et al., 2018), among which the phenomenon of popularity bias and the underrepresentation of niche artists — the so-called *long-tail* — is possibly one of the most studied (Celma, 2010). Nonetheless, several questions open the way to novel musically motivated analysis. What is the role of music recommendation diversity in shaping the listeners' experience? What are the implications of the emergence of diversity-related phenomena (such as filter bubbles, echo chambers, or balkanisation) on people's musical preferences? In the next section, reviewing the literature of diversity in Music RS we aim to understand at what stage the MIR research has contributed to this analysis.

3. Music Recommendation Diversity

Interactions between users and items are traditionally the main core of RS research, and Music RS does not differ in this aspect (Knees et al., 2019). Inspired by studies on the semiology of music (Nattiez and Dunsby, 1977; Molino et al., 1990), we can map the two elements of RS research, users and items, to two distinguishable domains, interdependent and both influential on the nature of music: the Poietic and the Esthetic domain. The *Poietic* domain (from Greek: *poiētikós*, 'creative') includes the works' creative processes, influenced by aspects such as the composer formation, musical theories, or the historical context in which the work is created. The *Esthetic* domain (from Greek: *aísthēsis*, 'perception') comprehends the aspects related to the listeners, hence their musical background, the historical situation, their perception of the musical work, and their musical knowledge. Reviewing the literature of Music RS research, we then present studies related to diversity mapping items and users to such domains (**Figure 1**).

3.1 Poietic domain – the item side

Music RS can be designed to recommend different categories of items, such as artists (e.g. Celma and Cano, 2008) or tracks (e.g. Kamehkhosh and Jannach, 2017). Several works in the RS diversity-related literature make use of Listening Event (LE) datasets, representing the interactions between items and users, for validating models and techniques by means of empirical analysis. In particular, data from the online music service Last.fm[4] is a widely used resource for many in the RS community (e.g. Vargas and Castells, 2011; Ribeiro et al., 2015; Kapoor et al., 2015; Ekstrand et al., 2018). Nonetheless, focusing on the MIR literature, we can have a more detailed understanding of how item diversity has been approached in the music domain.

A first line of research identifies diversity as the count of different items with which users interact, averaged and aggregated with different logics which often can be traced back to the dual-concept diversity. Schedl and Hauger (2015) use the number of listened-to tracks and their musical genres as a proxy to characterise users' musical taste in terms of diversity. A similar logic is proposed by Ferwerda and Schedl (2016), where diversity of listening behaviours is analyzed at a country-level. Again

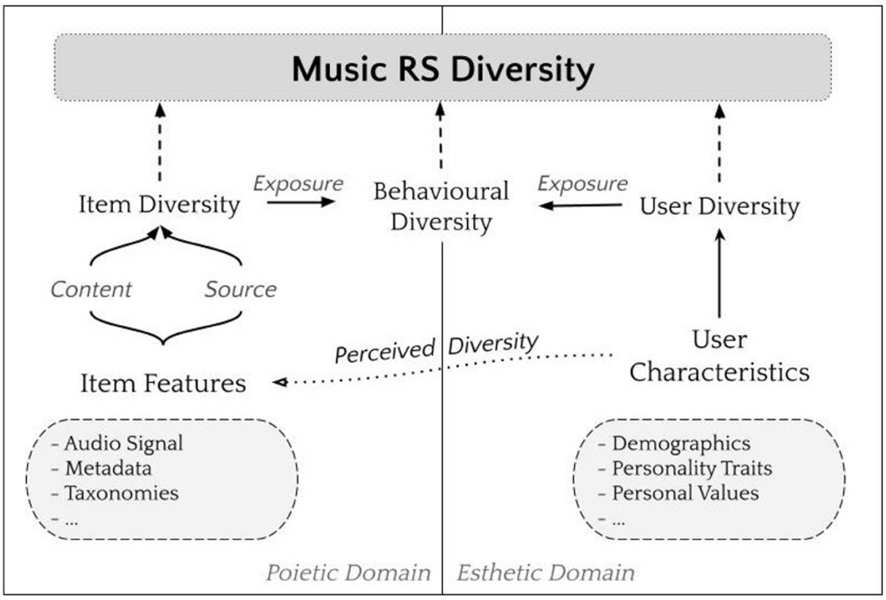

Figure 1: Mind map of elements constituting Music RS diversity. *Behavioural diversity*, for instance represented by listening events, is measured when users access the information provided by the items (*exposure*, Section 2.1). These connection points rely on one side on the *item diversity* (Section 3.1), built on content and source item features, and on the other side on *user diversity* (Section 3.2), with regards to their characteristics. Additionally, *perceived diversity* (Section 3.2.1) creates a bridge between the *Esthetic* and *Poietic* domains.

considering the country, Liu et al. (2017) measure the diversity by analyzing the distribution of artists' listening counts. The advantages of using such approaches rely on their not complex formulation and relatively simple implementation, and in addition they can be computed using only the listening events, eventually with artist or genre metadata. The main drawback is that they do not use any additional features to differentiate between items, risking to oversimplify the nature of concepts such as music genres (see Section 4.1).

A second line of work, building on top of the distribution of the user-item interactions, makes use of distance spaces containing additional information to diversify the items. For example, Park et al. (2015) and Way et al. (2019) use the Rao-Stirling diversity index, where a further dimension representing the closeness between items, genres in the former case and artists in the latter, is computed thanks to a co-consumption matrix. Porcaro and Gómez (2019) build an embedding space modeling user-generated tags to compute the diversity of a playlist. Similarly, Anderson et al. (2020) create a song-embedding space using user-generated playlists, from which a diversity score is derived. Although these data-driven approaches are able to compute items' fine-grained features for estimating diversity going beyond the dual-concept logic, it is also true that they are generally expensive in terms of data and computational resources.

Furthermore, approaches based on latent features extracted using matrix factorisation (MF) techniques have been proposed. For instance, Ferwerda et al. (2017a) and Robinson et al. (2020) compute diversity using the Euclidean distance between item vectors in the MF space. A great advantage of these methods is that they require only the user-item interaction matrix, however the little interpretability of the latent space makes it difficult to understand what are the item characteristics that determine the diversity. Alternative approaches using entropy-related metrics can be found, for instance using Shannon entropy (Ferwerda and Schedl, 2016; Zhou et al., 2018), and the Herfindahl-Hirschman index (Datta et al., 2018; Poulain and Tarissan, 2020).

What most of the aforementioned works share is their common perspective of measuring some sort of item diversity connected with the users' behaviours, focusing mostly on exposure diversity (**Table 2**). Content and source diversity instead are considered mainly in works

Table 2: List of works analyzing users' *behavioural diversity* in the music domain, presented in chronological order.

Reference	Diversity metric definition(s) Dataset(s)
Farrahi et al. (2014)	· Number of unique genres associated with the artists listened to by a user. MMTD (Hauger et al., 2013).
Schedl and Hauger (2015)	· Users' average track listening frequency; number of distinct track genres. Last.fm LEs.
Ferwerda et al. (2016b)	· Aggregation of each user's listening history by artist and genre. LFM-1b (Schedl, 2016).
Ferwerda and Schedl (2016)	· Overall volume of genre occurrences; relative listening volume exceeding one per mille; Shannon index computed over artist genre. LFM-1b (Schedl, 2016).
Park et al. (2015)	· Rao-Stirling index computed over artist genre. Last.fm users' top artists.
Datta et al. (2018)	· Log number of unique artists, songs, and genres listened to; number of unique top artists in a user's geographic region divided by the number of unique artists listened to over the same time period; Herfindahl index computed over a user's weekly plays. Spotify LEs.
Wang et al. (2018)	· Ratio of unique artists in a user's playlists over all the artists listened to by the user; same ratio computed over artist genre. Last.fm 1K (Celma, 2010).
Li et al. (2018)	· Hill-type true diversity (Rao-Stirling index) computed over album genre. Xiami LEs.
Way et al. (2019)	· Rao-Stirling index computed over artist genre. Spotify LEs.
Poulain and Tarissan (2020)	· Herfindahl-Hirschman index computed over tripartite graphs (users, tracks, tags). MSD (Bertin-Mahieux et al., 2011); Amazon Dataset (McAuley et al., 2015).
Anderson et al. (2020)	· Average cosine similarity between a track embedding and the average of the user's track embeddings. Spotify LEs.
Kowald et al. (2021)	· Cosine similarity computed over the users' track genre distributions. LFM-BeyMS (Kowald et al., 2021), subset of LFM-1b (Schedl, 2016).

centered on the analysis of music lists (e.g. playlists, recommendation lists, sessions), where however the user is often left aside (**Table 3**). Grouping users by their diversity is intended as grouping them by the diversity of the items they consumed, and in this behavioural perspective several important aspects related to the listener, the end-user of Music RS, are neglected as discussed in the next section.

3.2 Esthetic domain – the user side

Understanding music listeners is a hard problem, due to the multifaceted nature on one side of human behaviour, and on the other side of the act of listening to music. This problem was neglected by the MIR community in its early stages (Schedl et al., 2013; Lee and Cunningham, 2013), but more awareness of it is emerging recently (Knees et al., 2019). As a starting point, in line with Knees et al. (2019), it is important to separate *individual* from *collective* aspects, elements interconnected but subject to different praxes. We refer to collective aspects when referring to aspects of the music listening shared among people belonging to a specific group, be it built on ethnic, geographical, generational, or other criteria.

3.2.1 Individual aspects

Scholars in the field of psychology of music have addressed for decades the study of the role of listening to music in people's lives (Rentfrow, 2012). The several functions identified with this act, which juxtaposed bring out its ubiquity, are symptomatic of the adversities emerging while designing Music RS adaptable to different roles: *entertainment*, *identity formation*, *escapism*, *mood management*, *self-determination*, and *social differentiation*, just to mention a few (Schäfer et al. 2013).

Among the aspects characterising individuals' diversity needs while interacting with RS, personality traits have emerged as a focal feature for differentiating behaviours, largely investigated by the MIR community (Ferwerda et al., 2016a, 2017b; Jin et al., 2020; Lu and Tintarev, 2018). In these works, the five-factor model proposed by McCrae and John (1992) is a commonly accepted taxonomy which groups personality traits in five main dimensions: *Extraversion*, *Agreeableness*, *Conscientiousness*, *Neuroticism*, and *Openness to Experience*. Analyzing the correlation between these factors and music preferences, researchers have highlighted points of intersection between personal traits and the demand to diversify the listening experience. New directions have also been explored concerning the relationship between musical taste and personal values (Manolios et al., 2019). An important outcome of these studies is the differentiation between metric-based diversity, as measurement based on designed features extractable by algorithmic processes, and *perceived diversity*, hence how people evaluate a degree of diversity based on their personality, background, and beliefs (see Section 4.2).

3.2.2 Collective aspects

Acknowledging music as a social phenomenon entails the understanding of the overlap between individual practices and collective habits. Analyzing the interactions between people and musical objects, it is possible to observe a dual structure, where social groups are formed based on shared interests and taste, and parallelly genres and subcultures are dependent on their publics (DiMaggio, 2011). This intuition is at the core of several recommendation algorithms belonging to Collaborative Filtering (CF)

Table 3: List of works analysing *item diversity* in the music domain, presented in chronological order. We refer to Ziegler et al. (2005) for the formula of the Intra-List Diversity (ILD).

Reference	Diversity metric definition(s) Dataset(s)
Slaney and White (2006)	· Distribution of points in an 11-dimensional genre space computed over tracks' acoustic features. WebJay playlists.
Ferwerda et al. (2017a)	· ILD using Euclidean distance computed over the latent factor of item-user matrix factorisation. Last.fm LEs; LFM-1b (Schedl, 2016).
Lu and Tintarev (2018)	· ILD computed over weighted combinations of several diversity degrees for different attributes (release time, artist, genre, tempo, key). Spotify users' preferred songs; Echo Nest Taste Profile Subset (Bertin-Mahieux et al., 2011).
Porcaro and Gómez (2019)	· ILD using cosine distance computed over track tag embeddings. Art of the Mix playlists (Berenzweig et al., 2004); Yes.com radio playlists (Chen et al., 2012); MMTD (Hauger et al., 2013); Deezer users' playlists.
Knees and Hübler (2019)	· Simpson index computed over tracks' record labels. MPD (Chen et al., 2018).
Robinson et al. (2020)	· ILD using Euclidean distance computed over the latent factor of item-user matrix factorisation. Last.fm LEs.
Jin et al. (2020)	· ILD using Jaccard Index computed over track genre. Spotify users' recommendations.

methods, one of the widespread frameworks in the RS panorama (Ricci et al., 2015). The idea that *similar users like similar items* simplifies sociological aspects of group formations, but still is no stranger to social phenomena. Not surprisingly, and in line with Bourdieu's view that 'nothing more clearly affirms one's "class", nothing more infallibly classifies, than tastes in music' (Bourdieu, 1984), one among the first information filtering systems based on social information was a personalised Music RS named *Ringo* (Shardanand and Maes, 1995).

A huge limitation when studying collective aspects in Music RS research is the lack of data available to perform diversity analysis (see Section 4.3). Indeed, when characterising groups of listeners, often only country-related information can be exploited. Several examples of cross-country analysis can be found in the MIR literature. For instance, Ferwerda et al. (2016b) use Hofstede's cultural dimensions (Hofstede, 1991) to investigate users' diversity needs across countries. Liu et al. (2018), along with such dimensions, consider economic and linguistic diversity when modelling distances between users' country of origin. Alternatively, a characterisation of users' diversity based on socio-economic factors is presented by Park et al. (2015). Nonetheless, the use of the country's information as a proxy for classifying individuals can misrepresent the idea of culture with national culture, stigmatising aspects which however are not representative of multicultural environments (McSweeney, 2002). An alternative approach of using country information has been proposed by Schedl et al. (2021), where country archetypes are created based on listening preferences.

4. Challenges and Research Gaps

We have presented what so far have been the research directions in which diversity has been investigated in the Music RS field. Most of the work has focused on establishing methods to estimate the diversity of users' behaviours when interacting with recommendations, using as a proxy the diversity of the items consumed. Few works have also considered how recommendations interplay with the diversity of users' characteristics, whether at an individual level such as personality traits, or at a collective level such as country of origin. Nonetheless, interdisciplinary perspectives while designing Music RS are often overlooked, a trend already highlighted by Laplante (2014). This motivates us to provide three examples wherein undertaking an interdisciplinary attitude could help in designing diversity-aware Music RS systems.

4.1 Item diversity and music classification

An aspect to consider when dealing with the measurement of item diversity in music RS research is the long-standing debate about the classification of music and culture (Moles, 1967; Bourdieu, 1984; DiMaggio, 1987). Despite the dynamical, intrinsic, ambiguous, and context-dependent nature of concepts such as genre and style (Aucouturier and Pachet, 2003; Johansson, 2016), they have been historically used by the MIR community in different frameworks, mainly while presented in the form of tags (Lamere, 2008). However, when represented as tags, genres and styles are often deprived of meaningful historical and societal characteristics. Making such abstract concepts understandable by a machine is still an open question in MIR, and with current methods may still prove elusive, as observed by Sturm (2014) for the task of Music Genre Recognition. For addressing this challenge, alternative approaches to the use of a fixed taxonomy could be considered, as done by Vlegels and Lievens (2017) and Way et al. (2019), where the classification of items is based on the analysis of listeners' behaviour, enhancing the duality of cultural networks (DiMaggio, 2011).

4.2 User diversity and the musical Self

Contextualising the relationships between individual listening experience and Music RS, two facets of these technologies can be identified following the work done by Foucault (1988): on one side as *technologies of power*, influencing the behaviours of individuals, on the other as *technologies of the self*, providing a tool for transforming oneself. What are the expectations when receiving recommendations according to the image we have about our musical Self? Diversity here plays a key role because the urge to diversify can emerge simply by being exposed to such systems, affecting how we behave. As observed by a Last.fm user reflecting on her listening habits (Karakayali et al., 2018):

> Last.fm has changed me. Made me too self-conscious of my listening habits. Before, I'd play the same artist for days and days, but now I constantly struggle to diversify. I recently made a playlist called "diversify!"[...].

Under this lens, while designing Music RS to consider the dichotomy proposed by Roth (2019), where algorithmic influence on individuals' behaviours is classified into *read our mind* and *change our mind* processes, can lead to a deeper understanding of the interactions between listeners and Music RS.

4.3 User diversity and social background

Defining the relationships between social groups and musical tastes, Bourdieu's perspective in *Distinction*, and the so-called *omnivore thesis* by Peterson can be considered as two well-established theories evidencing underlying mechanisms of social interactions with cultural objects (Bourdieu, 1984; Peterson, 1992). Bourdieu's work focuses on the analysis of taste formation and definition in relation to social status, showing how economic, cultural, and social capitals play a central role in these processes. A decade later, Peterson presented the omnivore-univore model to describe audience segmentation in the USA during the early 1990's. In his model, omnivore refers to consumption habits of high-status participants characterised by a tendency to appreciate a wide variety of cultural products, while univore represents the part of the population used to consume few specific categories of cultural products.

The debate around such theories is still active (see Coulangeon and Lemel (2007); Atkinson (2011)), but both converge on the idea that the diversity of listening habits cannot be detached from the diversity of the social and cultural background in which people build their own experiences. In this respect, the work by Park et al. (2015) is a notable example of using socio-economic information to characterise the listeners' background, which goes beyond the cross-country analysis often pursued by MIR scholars.

5. Conclusions and Path Forward

Although we have some intuitions of how diversity can be treated when interacting with music recommendations, the lack of a bigger picture in which to frame such diversity analysis is a gap that we are witnessing (Porcaro et al. 2019). To imagine how diversity can be included as a design principle for the next generation of Music RS is an open debate for the MIR community, to which we aim to contribute with this overview. In the future, we foresee the definition of a set of practical indications to follow when approaching diversity in Music RS research.

As a starting point, we believe that it is critical for practitioners to explain how diversity is inferred while designing diversity-aware Music RS. It could be done by identifying which components are investigated (source, content, exposure), which perspectives are undertaken when designing algorithmic procedures (individual autonomy, deliberative, or adversarial), but also what impact, being it positive or negative, could we expect by the introduction of a specific design. This latter aspect is still underexplored in MIR research, and having a deeper understanding of the dynamics of the interaction between listeners and RS is without doubt a core issue. The increasing interest by the RS community in simulation-based frameworks and longitudinal studies paves the way to new findings which can be applied in the future design of Music RS (Ferraro et al., 2020). Furthermore, it is also important to investigate how different dimensions of diversity correlate when people interact with Music RS. Could we guarantee that the diversity of the items is not influential on the diversity of the users, and vice versa? The emergence of streaming services which target a music genre (e.g. IDAGIO[5] for classical music), or developed in a specific region of the world (e.g. Anghami[6] in the Middle East), poses new questions about how Music RS can be designed in scenarios wherein a globalist vision could fail in representing the peculiarities of artists and listeners.

In conclusion, we share the call for an interdisciplinary effort made by Born (2020), a necessary step to escape from the *technological solutionism* which has partly driven the Music RS research roadmap until now, a road which however may be full of traps (Selbst et al., 2019; Seaver, 2019).

Notes

1 We choose to follow this approach inspired by the work of Benjamin (2019, Chapter 5).
2 https://www.ifpi.org/ifpi-global-music-report-2019.
3 Under a multistakeholder perspective (Abdollahpouri et al., 2020), other actors influencing the diversity of the music listened to may be considered, for instance the platform providers, i.e., the music streaming services in the case of Music RS.
4 https://www.last.fm.
5 https://www.idagio.com.
6 https://www.anghami.com.

Acknowledgements

This work is partially supported by the European Commission under the TROMPA project (H2020 – grant agreement No. 770376).

This work is also partially supported by the HUMAINT programme (Human Behaviour and Machine Intelligence), Joint Research Centre, European Commission.

The project leading to these results received funding from "la Caixa" Foundation (ID 100010434), under the agreement LCF/PR/PR16/51110009.

Competing Interests

Emilia Gómez is a co-Editor-in-Chief of the Transactions of the International Society for Music Information Retrieval. She had no involvement in the review and editorial processing of this article. The authors have no other competing interests to declare.

References

Abdollahpouri, H., Adomavicius, G., Burke, R., Guy, I., Jannach, D., Kamishima, T., Krasnodebski, J., and **Pizzato, L.** (2020). Multistakeholder recommendation: Survey and research directions. *User Modeling and User-Adapted Interaction*, 30(1): 127–158. DOI: https://doi.org/10.1007/s11257-019-09256-1

Anderson, A., Maystre, L., Mehrotra, R., Anderson, I., and **Lalmas, M.** (2020). Algorithmic effects on the diversity of consumption on Spotify. In *Proceedings of The Web Conference 2020*, pages 2155–2165. DOI: https://doi.org/10.1145/3366423.3380281

Atkinson, W. (2011). The context and genesis of musical tastes: Omnivorousness debunked, Bourdieu buttressed. *Poetics*, 39(3): 169–186. DOI: https://doi.org/10.1016/j.poetic.2011.03.002

Aucouturier, J. J., and **Pachet, F.** (2003). Representing musical genre: A state of the art. *Journal of New Music Research*, 31(1): 83–93. DOI: https://doi.org/10.1076/jnmr.32.1.83.16801

Baeza-Yates, R. (2018). Bias on the web. *Communications of the ACM*, 61(6): 54–61. DOI: https://doi.org/10.1145/3209581

Barocas, S., and **Selbst, A. D.** (2014). Big data's disparate impact. *California Law Review*, 104(3): 671–732.

Benjamin, R. (2019). *Race After Technology*. Polity.

Berenzweig, A., Logan, B., Ellis, D. P., and **Whitman, B.** (2004). A large-scale evaluation of acoustic and subjective music-similarity measures. *Computer Music Journal*, 28(2): 63–76. DOI: https://doi.org/10.1162/014892604323112257

Bertin-Mahieux, T., Ellis, D. P. W., Whitman, B., and **Lamere, P.** (2011). The Million Song Dataset. In *Proceedings of the 12th International Society for Music Information Retrieval Conference*, pages 591–596.

Born, G. (2020). Diversifying MIR: Knowledge and real-world challenges, and new interdisciplinary futures. *Transactions of the International Society for Music Information Retrieval*, 3(1): 193–204. DOI: https://doi.org/10.5334/tismir.58

Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgement of Taste*. Routledge.

Bozdag, E., and **van den Hoven, J.** (2015). Breaking the filter bubble: Democracy and design. *Ethics and Information Technology*, 17(4): 249–265. DOI: https://doi.org/10.1007/s10676-015-9380-y

Castells, P., Hurley, N. J., and **Vargas, S.** (2015). Novelty and diversity in recommender systems. In Ricci, F., Rokach, L., and Shapira, B., editors, *Recommender Systems Handbook*, pages 881–918. Springer, Boston, MA. DOI: https://doi.org/10.1007/978-1-4899-7637-6_26

Celma, Ò. (2010). *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer-Verlag Berlin Heidelberg.

Celma, Ò., and **Cano, P.** (2008). From hits to niches? Or how popular artists can bias music recommendation and discovery. In *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*, pages 1–8. DOI: https://doi.org/10.1145/1722149.1722154

Chen, C.-W., Lamere, P., Schedl, M., and **Zamani, H.** (2018). RecSys Challenge 2018: Automatic music playlist continuation. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 527–528. DOI: https://doi.org/10.1145/3240323.3240342

Chen, S., Moore, J. L., Turnbull, D., and **Joachims, T.** (2012). Playlist prediction via metric embedding. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 714–722. DOI: https://doi.org/10.1145/2339530.2339643

Coulangeon, P., and **Lemel, Y.** (2007). Is 'distinction' really outdated? Questioning the meaning of the omnivorization of musical taste in contemporary France. *Poetics*, 35(2–3): 93–111. DOI: https://doi.org/10.1016/j.poetic.2007.03.006

Datta, H., Knox, G., and **Bronnenberg, B. J.** (2018). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. *Marketing Science*, 37(1): 5–21. DOI: https://doi.org/10.1287/mksc.2017.1051

DiMaggio, P. (1987). Classification in art. *American Sociological Review*, 52(4): 440–455. DOI: https://doi.org/10.2307/2095290

DiMaggio, P. (2011). Cultural networks. In Scott, J. and Carrington, P. J., editors, *The Sage Handbook of Social Network Analysis*, pages 286–310. SAGE Publications. DOI: https://doi.org/10.4135/9781446294413.n20

Drosou, M., Jagadish, H., Pitoura, E., and **Stoyanovich, J.** (2017). Diversity in big data: A review. *Big Data*, 5(2): 73–84. DOI: https://doi.org/10.1089/big.2016.0054

Ekstrand, M. D., Tian, M., Azpiazu, I. M., Ekstrand, J. D., Anuyah, O., McNeill, D., and **Pera, M. S.** (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*, pages 172–186.

Farrahi, K., Schedl, M., Vall, A., Hauger, D., and **Tkalčič, M.** (2014). Impact of listening behavior on music recommendation. In *Proceedings of the 15th International Society for Music Information Retrieval Conference*, pages 483–488.

Ferraro, A., Jannach, D., and **Serra, X.** (2020). Exploring longitudinal effects of session-based recommendations. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 474–479. DOI: https://doi.org/10.1145/3383313.3412213

Ferwerda, B., Graus, M., Vall, A., Tkalčič, M., and **Schedl, M.** (2016a). The influence of users' personality traits on satisfaction and attractiveness of diversified recommendation lists. In *Proceedings of the 4th Workshop on Emotions and Personality in Personalized Systems*, pages 43–47.

Ferwerda, B., Graus, M. P., Vall, A., Tkalčič, M., and **Schedl, M.** (2017a). How item discovery enabled by diversity leads to increased recommendation list attractiveness. In *Proceedings of the ACM Symposium on Applied Computing*, pages 1693–1696. DOI: https://doi.org/10.1145/3019612.3019899

Ferwerda, B., and **Schedl, M.** (2016). Investigating the relationship between diversity in music consumption behavior and cultural dimensions: A crosscountry analysis. In *Proceedings of the 1stWorkshop on Surprise, Opposition, and Obstruction in Adaptive and Personalized Systems*.

Ferwerda, B., Tkalčič, M., and **Schedl, M.** (2017b). Personality traits and music genres: What do people prefer to listen to? In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, pages 285–288. DOI: https://doi.org/10.1145/3079628.3079693

Ferwerda, B., Vall, A., Tkalčič, M., and **Schedl, M.** (2016b). Exploring music diversity needs across countries. In *Proceedings of the 24th Conference on User Modeling Adaptation and Personalization*, pages 287–288. DOI: https://doi.org/10.1145/2930238.2930262

Foucault, M. (1988). *Technologies of the Self: A Seminar with Michel Foucault*. University of Massachusetts Press.

Freire, A., Porcaro, L., and **Gómez, E.** (2021). Measuring diversity of artificial intelligence conferences. In *Proceedings of 2nd Workshop on Diversity in Artificial Intelligence*, pages 39–50.

Friedman, B., Kahn, P. H., and **Borning, A.** (2013). Value sensitive design and information systems. In Doorn, N., Schuurbiers, D., van de Poel, I., and Gorman, M. E., editors, *Early Engagement and New Technologies: Opening up the Laboratory. Philosophy of Engineering and Technology*, volume 16. Springer, Dordrecht. DOI: https://doi.org/10.1007/978-94-007-7844-3_4

Friedman, B., and **Nissenbaum, H.** (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3): 330–347. DOI: https://doi.org/10.1145/230538.230561

Gómez, E., Charisi, V., Tolan, S., Miron, M., Martinez Plumed, F., and **Planas, E.** (2021). *HUMAINT: Understanding the Impact of Artificial Intelligence on Human Behaviour*. European Union, Publications Office of the European Union, Luxembourg.

Grenier, L. (1989). From diversity to difference: The case of socio-cultural studies of music. *New Formations*, 1989(9).

Hauger, D., Schedl, M., Košir, A., and **Tkalčič, M.** (2013). The Million Musical Tweets Dataset: What can we learn from microblogs. In *Proceedings of the 14th International Society for Music Information Retrieval Conference*, pages 189–194.

Helberger, N. (2011). Diversity by design. *Journal of Information Policy*, 1(2011): 441–469. DOI: https://doi.org/10.5325/jinfopoli.1.2011.0441

Helberger, N., Karppinen, K., and **D'Acunto, L.** (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication and Society*, 21(2): 191–207. DOI: https://doi.org/10.1080/1369118X.2016.1271900

Hofstede, G. (1991). *Cultures and Organizations: Software of the Mind*. McGraw-Hill Book Company.

Holzapfel, A., Sturm, B. L., and **Coeckelbergh, M.** (2018). Ethical dimensions of music information retrieval technology. *Transactions of the International Society for Music Information Retrieval*, 1(1): 44–55. DOI: https://doi.org/10.5334/tismir.13

Huron, D. (2004). Issues and prospects in studying cognitive cultural diversity. In *Proceedings of the 8th International Conference on Music Perception and Cognition*, pages 93–96.

Jannach, D., and **Bauer, C.** (2020). Escaping the Mcnamara Fallacy: Toward more impactful recommender systems research. *AI Magazine*, 41(4): 79–95. DOI: https://doi.org/10.1609/aimag.v41i4.5312

Jin, Y., Tintarev, N., Htun, N. N., and **Verbert, K.** (2020). Effects of personal characteristics in control-oriented user interfaces for music recommender systems. *User Modeling and User-Adapted Interaction*, 30(2): 199–249. DOI: https://doi.org/10.1007/s11257-019-09247-2

Johansson, M. S. (2016). Making sense of genre and style in the age of transcultural reproduction. *International Review of the Aesthetics and Sociology of Music*, 47(1): 45–62.

Kamehkhosh, I., and **Jannach, D.** (2017). User perception of next-track music recommendations. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, pages 113–121. DOI: https://doi.org/10.1145/3079628.3079668

Kaminskas, M., and **Bridge, D.** (2016). Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-accuracy objectives in recommender systems. *ACM Transactions on Interactive Intelligent Systems*, 7(1): 1–42. DOI: https://doi.org/10.1145/2926720

Kapoor, K., Kumar, V., Terveen, L., Konstan, J. A., and **Schrater, P.** (2015). "I like to explore sometimes": Adapting to dynamic user novelty preferences. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 19–26. DOI: https://doi.org/10.1145/2792838.2800172

Karakayali, N., Kostem, B., and **Galip, I.** (2018). Recommendation systems as technologies of the self: Algorithmic control and the formation of music taste. *Theory, Culture and Society*, 35(2): 3–24. DOI: https://doi.org/10.1177/0263276417722391

Knees, P., and **Hübler, M.** (2019). Towards uncovering dataset biases: Investigating record label diversity in music playlists. In *Proceedings of the 1st Workshop on Designing Human-Centric MIR Systems*, pages 19–23.

Knees, P., Schedl, M., Ferwerda, B., and **Laplante, A.** (2019). User awareness in music recommender systems. In Augstein, M., Herder, E., and Wörndl, W., editors, *Personalized Human-Computer Interaction*, pages 223–252. De Gruyter Oldenbourg. DOI: https://doi.org/10.1515/9783110552485-009

Kowald, D., Muellner, P., Zangerle, E., Bauer, C., Schedl, M., and **Lex, E.** (2021). Support the underground: Characteristics of beyond-mainstream music listeners. *EPJ Data Science*, 10(1): 14. DOI: https://doi.org/10.1140/epjds/s13688-021-00268-9

Kunaver, M., and **Požrl, T.** (2017). Diversity in recommender systems: A survey. *Knowledge-Based Systems*, 123: 154–162. DOI: https://doi.org/10.1016/j.knosys.2017.02.009

Lamere, P. (2008). Social tagging and music information retrieval. *Journal of New Music Research*, 37(2): 101–114. DOI: https://doi.org/10.1080/09298210802479284

Laplante, A. (2014). Improving music recommender systems: What can we learn from research on music tastes? In *Proceedings of the 15th International Society for Music Information Retrieval Conference*, pages 451–456.

Lee, J. H., and **Cunningham, S. J.** (2013). Toward an understanding of the history and impact of user studies in music information retrieval. *Journal of Intelligent Information Systems*, 41(3): 499–521. DOI: https://doi.org/10.1007/s10844-013-0259-2

Li, H., Han, X. P., Lü, L., and **Pan, Z.** (2018). Measuring diversity of music tastes in online musical society. *International Journal of Modern Physics C*, 29(5): 1–10. DOI: https://doi.org/10.1142/S0129183118400065

Liebman, E., and **Stone, P.** (2020). Artificial musical intelligence: A survey. *Computing Research Repository*, pages 1–99.

Liu, M., Hu, X., and **Schedl, M.** (2017). Artist preferences and cultural, socio-economic distances across countries: A big data perspective. In *Proceedings of the 18th International Society for Music Information Retrieval Conference*, pages 103–111.

Liu, M., Hu, X., and **Schedl, M.** (2018). The relation of culture, socio-economics, and friendship to music preferences: A large-scale, cross-country study. *PLoS ONE*, 13(12): 1–29. DOI: https://doi.org/10.1371/journal.pone.0208186

Loecherbach, F., Moeller, J., Trilling, D., and **van Atteveldt, W.** (2020). The unified framework of media diversity: A systematic literature review. *Digital*

Journalism, 8(5): 605–642. DOI: https://doi.org/10.1080/21670811.2020.1764374

Lu, F., and **Tintarev, N.** (2018). A diversity adjusting strategy with personality for music recommendation. In *Proceedings of the 5th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems*, pages 7–14.

Lunardi, G. M. (2019). Representing the filter bubble: Towards a model to diversification in news. In Guizzardi, G., Gailly, F., and Suzana Pitangueira Maciel, R., editors, *Advances in Conceptual Modeling*, pages 239–246. DOI: https://doi.org/10.1007/978-3-030-34146-6_22

Manolios, S., Hanjalic, A., and **Liem, C. C.** (2019). The influence of personal values on music taste: Towards value-based music recommendations. *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 501–505. DOI: https://doi.org/10.1145/3298689.3347021

McAuley, J., Targett, C., Shi, Q., and **van den Hengel, A.** (2015). Image-based recommendations on styles and substitutes. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 43–52. DOI: https://doi.org/10.1145/2766462.2767755

McCrae, R. R., and **John, O. P.** (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2): 175–215. DOI: https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McDonald, D. G., and **Dimmick, J.** (2003). The conceptualization and measurement of diversity. *Communication Research*, 30(1): 60–79. DOI: https://doi.org/10.1177/0093650202239026

McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith – a failure of analysis. *Human Relations*, 55(1): 89–118. DOI: https://doi.org/10.1177/0018726702551004

Mitchell, M., Baker, D., Denton, E., Hutchinson, B., Hanna, A., and **Morgenstern, J.** (2020). Diversity and inclusion metrics in subset selection. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 117–123. DOI: https://doi.org/10.1145/3375627.3375832

Moles, A. A. (1967). *Sociodynamique de la Culture*. Paris, Mouton. DOI: https://doi.org/10.1515/9783111672403

Molino, J., Underwood, J. A., and **Ayrey, C.** (1990). Musical fact and the semiology of music. *Music Analysis*, 9(2): 105–156. DOI: https://doi.org/10.2307/854225

Napoli, P. M. (1999). Deconstructing the diversity principle. *Journal of Communication*, 49(4): 7–34. DOI: https://doi.org/10.1111/j.1460-2466.1999.tb02815.x

Nattiez, J.-J., and **Dunsby, J. M.** (1977). Fondements d'une sémiologie de la musique. *Perspectives of New Music*, 15(2): 226–233. DOI: https://doi.org/10.2307/832821

Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., and **Konstan, J. A.** (2014). Exploring the filter bubble: The effect of using recommender systems on content diversity. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 677–686. DOI: https://doi.org/10.1145/2566486.2568012

Olteanu, A., Castillo, C., Diaz, F., and **Kıcıman, E.** (2016). Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in Big Data*, 2: 1–47. DOI: https://doi.org/10.2139/ssrn.2886526

Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. The Penguin Press, New York, NY, USA.

Park, M., Weber, I., Naaman, M., and **Vieweg, S.** (2015). Understanding musical diversity via online social media. In *Proceedings of the 9th International AAAI Conference on Web and Social Media*, pages 308–317.

Peterson, R. A. (1992). Understanding audience segmentation: From elite and mass to omnivore and univore. *Poetics*, 21(4): 243–258. DOI: https://doi.org/10.1016/0304-422X(92)90008-Q

Porcaro, L., Castillo, C., and **Gómez, E.** (2019). Music recommendation diversity: A tentative framework and preliminary results. In *Proceedings of the 1st Workshop on Designing Human-Centric MIR Systems*, pages 11–15.

Porcaro, L., and **Gómez, E.** (2019). 20 years of playlists: A statistical analysis on popularity and diversity. In *Proceedings of the 20th International Society for Music Information Retrieval Conference*, pages 4–11.

Poulain, R., and **Tarissan, F.** (2020). Investigating the lack of diversity in user behavior: The case of musical content on online platforms. *Information Processing and Management*, 57(2): 1–18. DOI: https://doi.org/10.1016/j.ipm.2019.102169

Rentfrow, P. J. (2012). The role of music in everyday life: Current directions in the social psychology of music. *Social and Personality Psychology Compass*, 6(5): 402–416. DOI: https://doi.org/10.1111/j.1751-9004.2012.00434.x

Ribeiro, M. T., Ziviani, N., Moura, E. S. D., Hata, I., Lacerda, A., and **Veloso, A.** (2015). Multiobjective Pareto-efficient approaches for recommender systems. *ACM Transactions on Intelligent Systems and Technology*, 5(4): 1–20. DOI: https://doi.org/10.1145/2629350

Ricci, F., Rokach, L., and **Shapira, B.** (2015). *Recommender Systems Handbook*. Springer New York Heidelberg Dordrecht London, 2nd edition. DOI: https://doi.org/10.1007/978-1-4899-7637-6

Robinson, K., Brown, D., and **Schedl, M.** (2020). User insights on diversity in music recommendation lists. In *Proceedings of the 21st International Society for Music Information Retrieval Conference*, pages 446–453.

Roth, C. (2019). Algorithmic distortion of informational landscapes. *Intellectica*, 70(1): 97–118.

Salamon, J. (2019). What's broken in music informatics research? Three uncomfortable statements. In *Proceedings of the 36th International Conference on Machine Learning*, pages 2012–2014.

Schäfer, T., Sedlmeier, P., Städtler, C., and **Huron, D.** (2013). The psychological functions of music listening. *Frontiers in Psychology*, 4: 1–33. DOI: https://doi.org/10.3389/fpsyg.2013.00511

Schedl, M. (2016). The LFM-1b Dataset for music retrieval and recommendation. In *Proceedings of the*

2016 ACM International Conference on Multimedia Retrieval, pages 103–110. DOI: https://doi.org/10.1145/2911996.2912004

Schedl, M., Bauer, C., Reisinger, W., Kowald, D., and **Lex, E.** (2021). Listener modeling and contextaware music recommendation based on country archetypes. *Frontiers in Artificial Intelligence*, 3: 1–21. DOI: https://doi.org/10.3389/frai.2020.508725

Schedl, M., Flexer, A., and **Urbano, J.** (2013). The neglected user in music information retrieval research. *Journal of Intelligent Information Systems*, 41: 523–539. DOI: https://doi.org/10.1007/s10844-013-0247-6

Schedl, M., and **Hauger, D.** (2015). Tailoring music recommendations to users by considering diversity, mainstreaminess, and novelty. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 947–950. DOI: https://doi.org/10.1145/2766462.2767763

Schedl, M., Zamani, H., Chen, C.-W., Deldjoo, Y., and **Elahi, M.** (2018). Current challenges and visions in music recommender systems research. *International Journal of Multimedia Information Retrieval*, 7: 95–116. DOI: https://doi.org/10.1007/s13735-018-0154-2

Seaver, N. (2019). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, 24(4): 421–436. DOI: https://doi.org/10.1177/1359183518820366

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., and **Vertesi, J.** (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, pages 59–68. DOI: https://doi.org/10.1145/3287560.3287598

Serra, X. (2011). A multicultural approach in music information research. In *Proceedings of the 12th International Society for Music Information Retrieval Conference*, pages 151–156.

Serra, X., Magas, M., Benetos, E., Chudy, M., Dixon, S., Flexer, A., Gómez, E., Gouyon, F., Herrera, P., Jordà, S., Paytuvi, O., Peeters, G., Schlüter, J., Vinet, H., and **Widmer, G.** (2013). Roadmap for music information research. http://www.mires.cc/files/MIRES_Roadmap_ver_1.0.0.pdf; accessed 21 October 2021.

Shardanand, U., and **Maes, P.** (1995). Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 210–217. DOI: https://doi.org/10.1145/223904.223931

Slaney, M., and **White, W.** (2006). Measuring playlist diversity for recommendation systems. In *Proceedings of the 1st ACM Workshop on Audio and Music Computing Multimedia*, pages 77–82. DOI: https://doi.org/10.1145/1178723.1178735

Steel, D., Fazelpour, S., Gillette, K., Crewe, B., and **Burgess, M.** (2018). Multiple diversity concepts and their ethical-epistemic implications. *European Journal for Philosophy of Science*, 8: 761–780. DOI: https://doi.org/10.1007/s13194-018-0209-5

Stirling, A. (2007). A general framework for analyzing diversity in science, technology and society. *Journal of The Royal Society Interface*, 4(15): 707–719. DOI: https://doi.org/10.1098/rsif.2007.0213

Sturm, B. L. (2014). A survey of evaluation in music genre recognition. In Nürnberger, A., Stober, S., Larsen, B., and Detyniecki, M., editors, *Adaptive Multimedia Retrieval: Semantics, Context, and Adaptation*, pages 29–66. Springer, Cham.

Sunstein, C. (2001). *Echo Chambers: Bush v. Gore Impeachment, and Beyond*. Princeton University Press.

UNESCO. (2001). UNESCO Universal Declaration on Cultural Diversity. http://portal.unesco.org/en/ev.php-URL_ID=13179&URL_DO=DO_TOPIC&URL_SECTION=201.html; accessed 15 October 2021.

Van Alstyne, M., and **Brynjolfsson, E.** (2005). Global village or cyber-Balkans? Modeling and measuring the integration of electronic communities. *Management Science*, 51(6): 851–868. DOI: https://doi.org/10.1287/mnsc.1050.0363

Vargas, S., and **Castells, P.** (2011). Rank and relevance in novelty and diversity metrics for recommender systems. In *Proceedings of the 5th ACM Conference on Recommender Systems*, pages 109–116. DOI: https://doi.org/10.1145/2043932.2043955

Vlegels, J., and **Lievens, J.** (2017). Music classification, genres, and taste patterns: A ground-up network analysis on the clustering of artist preferences. *Poetics*, 60: 76–89. DOI: https://doi.org/10.1016/j.poetic.2016.08.004

Wagner, E., and **Veloso, L.** (2019). Arts education and diversity: Terms and concepts. In Ferro, L., Wagner, E., Veloso, L., IJdens, T., and Teixeira Lopes, J., editors, *Arts and Cultural Education in a World of Diversity: ENO Yearbook 1*, pages 1–10. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-030-06007-7_14

Wang, M., Xiao, Y., Zheng, W., and **Jiao, X.** (2018). RNDM: A random walk method for music recommendation by considering novelty, diversity, and mainstream. In *Proceedings of the IEEE 30th International Conference on Tools with Artificial Intelligence*, pages 177–183. DOI: https://doi.org/10.1109/ICTAI.2018.00036

Way, S. F., Gil, S., Anderson, I., and **Clauset, A.** (2019). Environmental changes and the dynamics of musical identity. In *Proceedings of the 13th International AAAI Conference on Web and Social Media*, pages 527–536.

West, S. M., Whittaker, M., and **Crawford, K.** (2019). *Discriminating Systems: Gender, Race and Power in AI*. AI Now Institute.

Zhou, Z., Xu, K., and **Zhao, J.** (2018). Homophily of music listening in online social networks of China. *Social Networks*, 55: 160–169. DOI: https://doi.org/10.1016/j.socnet.2018.07.001

Ziegler, C.-N., McNee, S. M., Konstan, J. A., and **Lausen, G.** (2005). Improving recommendation lists through topic diversification. In *Proceedings of the 14th International Conference on World Wide Web*, pages 22–32. DOI: https://doi.org/10.1145/1060745.1060754

How to cite this article: Porcaro, L., Castillo, C., and Gómez, E. (2021). Diversity by Design in Music Recommender Systems. *Transactions of the International Society for Music Information Retrieval,* 4(1), pp. 114–126. DOI: https://doi.org/10.5334/tismir.106

Submitted: 18 March 2021 **Accepted:** 19 July 2021 **Published:** 02 November 2021

]u[*Transactions of the International Society for Music Information Retrieval* is a peer-reviewed open access journal published by Ubiquity Press. OPEN ACCESS