# INTELLIGENT USER INTERFACES FOR MUSIC DISCOVERY: THE PAST 20 YEARS AND WHAT'S TO COME

**Peter Knees** [1]     **Markus Schedl** [2]     **Masataka Goto** [3]

[1] Faculty of Informatics, TU Wien, Vienna, Austria

[2] Institute of Computational Perception, Johannes Kepler University Linz, Austria

[3] National Institute of Advanced Industrial Science and Technology (AIST), Japan

`peter.knees@tuwien.ac.at, markus.schedl@jku.at, m.goto@aist.go.jp`

## ABSTRACT

Providing means to assist the user in finding music is one of the original motivations underlying the research field known as Music Information Retrieval (MIR). Therefore, already the first edition of ISMIR in the year 2000 called for papers addressing the topic of "User interfaces for music IR". Since then, the way humans interact with technology to access and listen to music has substantially changed, not least driven by the advances of MIR and related research fields such as machine learning and recommender systems.

In this paper, we reflect on the evolution of MIR-driven user interfaces for music browsing and discovery over the past two decades. We argue that three major developments have transformed and shaped user interfaces during this period, each connected to a phase of new listening practices: first, connected to personal music collections, intelligent audio processing and content description algorithms that facilitate the automatic organization of repositories and finding music according to sound qualities; second, connected to collective web platforms, the exploitation of user-generated metadata pertaining to semantic descriptions; and third, connected to streaming services, the collection of online music interaction traces on a large scale and their exploitation in recommender systems.

We review and contextualize work from ISMIR and related venues from all three phases and extrapolate current developments to outline possible scenarios of music recommendation and listening interfaces of the future.

## 1. INTRODUCTION

With a history of five years of ISMIR conferences, in 2004, Downie [11] attempts to define the research field of Music Information Retrieval (MIR) as "a multidisciplinary research endeavor that strives to develop innovative content-based searching schemes, novel interfaces, and evolving networked delivery mechanisms in an effort to make the world's vast store of music accessible to all". Given the music industry landscape and how people listen to music 15 years later, this definition has not only stood the test of time, but also proven to be visionary.

With its origins in Information Retrieval research (cf. [5]), one of the original motivations underlying MIR was indeed to develop technology and provide means to assist the user in finding music. As the way humans interact with technology to access and listen to music has substantially changed since then, user interfaces for music discovery remain to be a pivotal element in MIR research. [1]

In this paper, we reflect on the evolution of MIR-driven user interfaces for music browsing and discovery over the past two decades—from organizing personal music collections to streaming a personalized selection from "the world's vast store of music". Therefore, we connect major developments that have transformed and shaped MIR research in general and user interfaces in particular to prevalent and emerging listening practices at the time. We identify three main phases that have each laid the foundation for the next and review work that focuses on the specific aspects of these phases.

First, we investigate the phase of growing digital personal music collections and interfaces built upon intelligent audio processing and content description algorithms in section 2. These algorithms facilitate the automatic organization of repositories and finding music in personal collections, as well as commercial repositories according to sound qualities. Second, in section 3, we investigate the emergence of collective web platforms and their exploitation for listening interfaces. The extracted user-generated metadata often pertains to semantic descriptions and complements the content-based methods that facilitated the developments of the preceding phase. This phase also constitutes an intermediate step towards exploitation of collective listening data, which is the driving force behind the third, and ongoing phase, which is connected to streaming services (section 4). Here, the collection of online music interaction traces on a large scale and their exploitation in recommender systems are defining elements.

---

[1] So do interfaces for "active listening," aiming at increasingly engaging the listener through augmented experiences and allowing also musically untrained listeners to gain deeper insights into aspects of the music they are consuming [20]. However, in this work we will not emphasize these interfaces.

Extrapolating these and other ongoing developments, we outline possible scenarios of music recommendation and listening interfaces of the future in section 5.

## 2. PHASE 1: CONTENT-BASED MUSIC RETRIEVAL INTERFACES

Based on the technological advancements in encoding and compression of audio signals (most notably mp3) together with the establishment of the Internet as mainstream communication medium and distribution channel, and, in rapid succession, the development of high capacity portable music players in the late 1990s, digital music has not only stirred up the IT industry, but also initiated a profound change in the way people "use" music.

At the time, the most popular and conventional interfaces for such music access display the list of bibliographic information (metadata) such as titles and artist names. When the number of musical pieces in a personal music collection was not large, music interfaces with the title list and mere text searches based on bibliographic information were useful enough to browse the whole collection to choose pieces to listen to. However, as the accessible collection grows, such simple interfaces become insufficient, and new research approaches targeting the retrieval, classification, and organization of music emerge.

"Intelligent" interfaces for music retrieval became a research field of interest with the developments in content-based music retrieval [6]. A landmark in this regard was the development of query by humming systems [31] and search engines indexing sound properties of loudness, pitch, and timbre [77] that initiated the emancipation of music search systems from traditional text- and metadata-based indexing and query interfaces. While interfaces were still very much targeted at presenting results in sequential order according to relevance to a query, in the early 2000s, MIR research proposed several alternatives to facilitate music discovery.

### 2.1 Map-based music browsing and discovery

Interfaces that allow content-based searches for music retrieval are useful when people can formulate good queries and especially when users are looking for a particular work, but sometimes it is difficult to come up with an appropriate query when faced with a huge music collection and vague search criteria. Interfaces for music browsing and discovery are therefore proposed to let users encounter unexpected but interesting musical pieces or artists. Visualization of a music collection is one way to provide users with various bird's-eye views and comprehensive interactions. The most popular visualization is to project musical pieces or artists onto a 2D or 3D space by using music similarity. 2D visualizations also lend themselves to being applied on tabletop interfaces for intuitive access and interaction, e.g. [30]. The trend of spatially arranging collections for exploration can be seen throughout the past 20 years and is still unbroken.
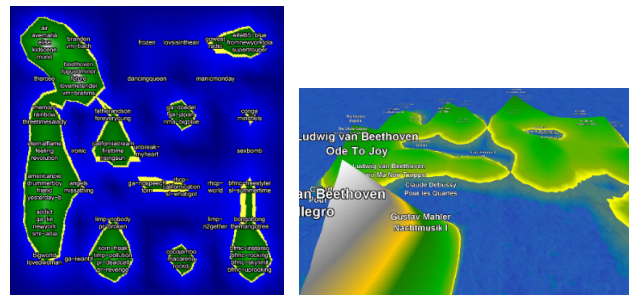


**Figure 1**. Islands of Music (left) and nepTune (right): music browsing interfaces that let a user explore a music collection by using a metaphor of "islands" visualizing self-organized clusters.

One of the earliest interfaces is *GenreSpace* [69] that visualizes musical pieces with genre-specific colors in a 3D space. Coloring of each piece is determined by automatic genre classification. The layout of pieces is determined by principal component analysis (PCA), which projects high-dimensional audio feature vectors into 3D positions. This idea is frequently used in other more recent interfaces.

Another early interface called *Islands of Music* [48] visualizes musical pieces on a 2D space representing an artificial landscape. It uses a self-organizing map (SOM) to arrange musical pieces so that similar pieces are located near each other. As shown on the left side in Figure 1, it uses a metaphor of "islands" that represent self-organized clusters of similar pieces. The denser the regions (more items in the same cluster), the higher the landscape (up to "mountains" for very dense regions). Sparse regions are represented by the ocean. This interface provides three different views corresponding to similarities based on three aspects: (1) timbre analysis, (2) rhythm analysis, and (3) metadata like artist and genre. A user can smoothly change focus from one view to another while exploring how the organization changes. Several extensions of the Islands of Music idea were proposed in following years. An aligned SOM is used by Pampalk et al. in [48] to enable a seamless shift of focus between clusterings created for different musical aspects, for instance, between a SOM created only on rhythm features and one created only on timbre features. The *nepTune* interface presented by Knees et al. in [36], as shown on the right side in Figure 1, enables exploration of music collections by navigating through a three-dimensional artificial landscape. Variants include a mobile version [27] and a larger-scale version using a growing hierarchical self-organizing map (GHSOM) [10] that automatically structures the map into hierarchically linked individual SOMs [57]. Neumayer et al. propose a method to automatically generate playlists by drawing a curve on the SOM visualization [47]. Lübbers and Jarke [42] present a browser employing multi-dimensional scaling (MDS) and SOM to create 3-dimensional landscapes. In contrast to the *Islands of Music* metaphor, they use an inverse height map, meaning that agglomerations of songs are visualized as valleys, while clusters are separated by mountains. Their interface further enables the user to adapt
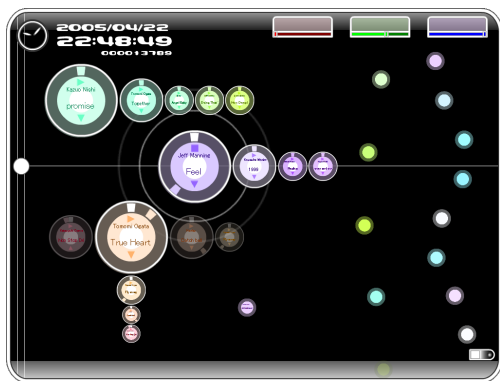
**Figure 2**. Musicream: A music playback interface that lets a user unexpectedly come across various musical pieces similar to those liked by the user.



**Figure 3**. "Reinventing the Wheel" (left) and its application on an iPod using the dial for browsing the collection (right).

the landscape by building or removing mountains, which triggers an adaptation of the underlying similarity measure.

Another SOM-based browsing interface is *Globe of Music* [40], which maps songs to a sphere instead of a plane by means of a GeoSOM [78]. Mörchen et al. [46] employ an emergent SOM and the U-map visualization technique [70] to color-code similarities between neighboring clusters. Vembu and Baumann incorporate a dictionary of musically related terms to describe similar artists [74].

In the *Search Inside the Music* application [38], Lamere and Eck use a three-dimensional MDS projection. Their interface provides different views that arrange images of album covers according to the output of the MDS, either in a cloud, a grid, or a spiral. Vad et al. [71] apply t-SNE [72] to mood- and emotion-related descriptors, which they infer from low-level acoustic features. The result of the data projection is visualized on a two-dimensional map, around which the authors build an interface to support the creation of playlists by drawing a path and by area selection.

While the above interfaces focus on musical pieces, interfaces focusing on artists have also been investigated. For example, *Artist Map* [73] is an interface that enables users to explore and discover artists. This interface projects artists onto a 2D space and visualizes them as small dots with genre-specific, tempo-specific, or year-specific colors. This visualization can also be used to create playlists by drawing paths and specifying regions.

Other examples use, e.g., metaphors of a "galaxy" or "cosmos," or extend visualizations with additional information. *MusicGalaxy* [65], for example, is an exploration interface that uses a similarity-preserving projection of musical pieces onto a 2D galaxy space. It takes timbre, rhythm, dynamics, and lyrics into account in computing the similarity and uses an adaptive non-linear multi-focus zoom lens that can simultaneously zoom multiple regions of interest while most interfaces support only a single region zooming. As a similar metaphor, "planetarium" has been used in *Songrium*,[2] a public web service for interactive visualization and exploration of *web-native music* on video sharing services [23]. It uses similarity-preserving

projection of pieces onto both 2D and 3D galaxy spaces and provides various functions: analysis and visualization of derivative works, and interactive chronological visualization and playback of musical pieces. *Instrudive* [66] enables users to browse and listen to musical pieces by focusing on instrumentation detected automatically. It visualizes each musical piece as a multicolored pie chart in which different colors denote different instruments. The ratios of the colors indicate relative duration in which the corresponding instruments appear in the piece.

## 2.2 Content-based filtering and sequential play

When a collection of music becomes huge, it is not feasible to visualize all pieces in the collection. Other types of interfaces that visualize a part of the music collection instead of the whole have also been proposed. An example is *Musicream* [21], a user interface that focuses on inducing active user interactions to discover and manage music in a huge collection. The idea behind Musicream is to see if people can break free from stereotyped thinking that music playback interfaces must be based on lists of song titles and artist names. To satisfy the desire "I want to hear something," it allows a user to unexpectedly come across various pieces similar to ones that the user likes. As shown on the right side in Figure 2, disk icons representing pieces flow one after another from top to bottom, and a user can select a disk and listen to it. By dragging a favorite disk in the flow, which serves as the query, the user can easily pick out other pieces similar to the query disk (attach similar disks) by using content-based similarity. In addition, to satisfy a desire like "I want to hear something my way," Musicream gives a user greater freedom of editing playlists by generating a playlist of playlists. Since all operations are automatically recorded, the user can also visit and retrieve a past state as if using a time machine.

The *FM4 Soundpark Player* makes content-based suggestions by showing up to five similar tracks in a graph-like manner [17] and constructing "mixtapes" from given start and end tracks [15]. *VocalFinder* [16] enables content-based retrieval of songs with vocals that have similar vocal timbre to the query song.

Visualization of a music collection is not always necessary to develop music interfaces. Stewart et al. [64] present

---

[2] https://songrium.jp

an interface that uses only sound auralization and haptic feedback to explore a large music collection in a two or three-dimensional space.

The article *"Reinventing the Wheel"* [51] revealed that a single-dial browsing device can be a useful interface for musical pieces stored on mobile music players. The whole collection is ordered in a circular locally-consistent playlist by using the Traveling Salesman algorithm so that similar pieces can be arranged adjacently. The user may simply turn the wheel to access different pieces. This interface also has the advantage of combining two different similarity measures, one based on timbre analysis and the other based on community metadata analysis. Figure 3 shows the conceptual prototype as well as an extended implementation on an Apple iPod [60], the most popular mobile listening device at the time.

## 3. PHASE 2: COLLABORATIVE AND AUTOMATIC SEMANTIC DESCRIPTION

While content-based analysis allowed for unprecedented views on music collections based on sound, interfaces built solely upon the extracted information were not able to "explain" the music contained or give semantically meaningful support for orientation within the collections. That is, while they are able to capture qualities of the sound of the contained music, they largely neglect human concepts of music organization, such as (sub-)genres or listening purposes, e.g., during activities like workouts. This information is however typically found on the web and ranges from user-generated tags to unstructured bits of expressed opinions (e.g., forum posts or comments in social media) to more detailed reviews and encyclopedic articles (containing, e.g., biographies and discography release histories). In MIR, this type of data is often referred to as community metadata or music context data [35]. These online "collaborative efforts" of describing music are resulting in a rich vocabulary of semantic labels and have shaped music retrieval interfaces towards music information systems. In parallel, platforms like *Last.fm* [3] (in this context better: *AudioScrobbler*), take advantage of users being increasingly always connected to the Internet and tracking listening events for the sake of identifying listening patterns and making recommendations, leading to the phase of automatic playlisting and music recommendation (cf. section 4). In this section, we focus on tags as a main driver of MIR research and music interfaces.

### 3.1 Collaborative platforms and music information systems

With music related information being ubiquitous on the web, dedicated web platforms that provide background knowledge on artists emerge, e.g., the AllMusic Guide, depending on editorial content. Using new technologies, such music information systems can, however, also be built by aggregating information extracted from various sources using text mining methods [59] or by taking advantage of
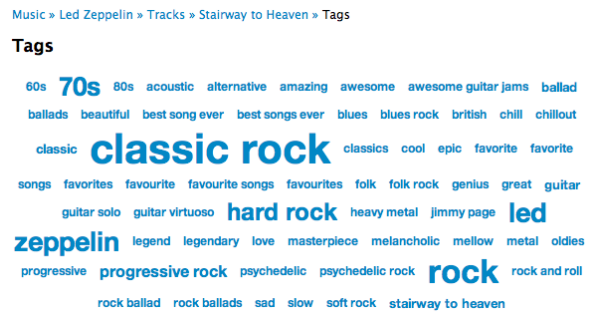
**Figure 4**. Last.fm tags of Led Zeppelin (source: https://musicmachinery.com/tag/lastfm/)

the "wisdom of the crowd" and building collaborative platform like the above mentioned Last.fm.

A central feature of Last.fm is to allow users to tag their music, ideally resulting in a democratic ground truth of what could be considered the semantic dimensions of the corresponding tracks, cf. Figure 4. However, typical problems arising with this type of information are noisy and non-trustworthy information as well as data sparsity and cold start issues mostly due to popularity biases, cf. [37].

MIR research during this phase has therefore dealt extensively with auto-tagging, i.e., automatically inferring semantic labels from the audio signal of a music piece (or related data), to overcome this shortcoming, e.g., [3, 12, 33, 44, 63, 68, 76].

Alternative approaches to generate semantic labels include human workforce. *TagATune* [39] is a game that pairs players across the Internet who try to determine whether they are listening to the same song by typing tags. In return for entertaining users, TagATune has collected interesting tags for a database of songs. Other examples of interfaces that were designed to collect useful information while engaging with music are *MajorMiner* [43], *HerdIt* [2], and *Moodswings* [34] (cf. section 4.1).

A more traditional way to obtain musically informed labels is to have human experts, e.g. trained musicians, manually label music tracks according to predefined musical categories. This approach is followed by the Music Genome Project, [4] and serves as the foundation of *Pandora's* automatic radio stations (cf. section 4).

As a consequence of these efforts, during this phase, the question of how to present and integrate this information into interfaces was secondary to the question of how to obtain it, as will become obvious next.

### 3.2 Visual interfaces

With the trend towards web-based interfaces, visualization and map based interfaces integrating semantic information have been proposed.

*MusicRainbow* [49] is a user interface for discovering unknown artists, which follows the above idea of a single-dial browsing device but features informative visualization. As shown in Figure 5, artists are mapped on a circular rainbow where colors represent different styles of music.
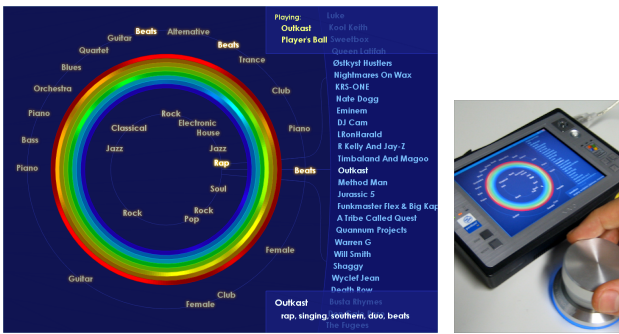
**Figure 5**. MusicRainbow: An artist discovery interface that enables a user to actively browse a music collection by using audio-based similarity and web-based labeling.

Similar artists are automatically mapped near each other by using the Traveling Salesman algorithm and summarized with word labels extracted from artist-related web pages. A user can rotate the rainbow by turning a knob and find an interesting artist by referring to the word labels.

The *nepTune* interface shown in Figure 1 also provides a mode that integrates text-based information extracted from artist web pages for supporting navigation in the 3D environment. To this end, labels referring to genres, instruments, origins, and eras serve as landmarks.

Other approaches explore music context data to visualize music over real geographical maps, rather than computing a clustering based on audio descriptors. For instance, Govaerts and Duval extract geographical information from biographies and integrate it into a visualization of radio station playlists [22]. Hauger and Schedl extract listening events and location information from microblogs and visualize both on a world map [25].

Lyrics are also important elements of music. By using semantic topics automatically estimated from lyrics, new types of visual interfaces for lyrics retrieval can be achieved. *LyricsRadar* [55] is a lyrics retrieval interface that uses latent Dirichlet allocation (LDA) to analyze topics of lyrics and visualizes the topic ratio for each song by using the topic radar chart. It then enables a user to find her favorite lyrics interactively. *Lyric Jumper* [67] is a lyrics-based music exploratory web service that enables a user to choose an artist based on topics of lyrics and find unfamiliar artists who have similar profile to her favorite artist. It uses an advanced topic model that incorporates an artist's profile of lyrics topics and provides various functions such as topic tendency visualization, artist ranking, artist recommendation, and lyric phrase recommendation.

## 4. PHASE 3: RECOMMENDER INTERFACES AND CONTINUOUS STREAMING

With ubiquitous Internet connection and a development of computer and entertainment systems to be always online, personal music collections have lost relevance to many people, as virtually all music content is available at all times. In essence, such subscription streaming services like Spotify, Pandora, Deezer, Amazon Music, or Apple

Music have transformed the music business and music listening alike.

Central element to these services is the aspect of personalization, i.e., providing foremost a user-tailored view onto the available collections of allegedly tens of millions of songs. Discovery of music is therefore performed by the system, based on the user profile of past interactions, rather than by the user herself.

Interfaces for music recommendation can support music listening in more personalized ways. Music recommendation typically models personal preferences of users by using their listening histories or explicit user feedback, e.g. [7, 62]. It then generates a set of recommended musical pieces or artists for each user. This recommendation can be implemented by using collaborative filtering based on users' past behaviors and exhibits patterns of music similarity not captured by content-based approaches [61]. When the playback order of recommended pieces is important, automatic playlist generation is also used, e.g. [4, 24, 45].

The main challenges of this type of algorithms are, as in all other domains of recommender systems, cold start problems. The approach taken to remedy these are again to integrate additional information on the music items to recommend, i.e. facets of content and metadata as applied in the earlier phases, by building hybrid recommenders on top of pure collaborative filtering. Additionally, context-awareness plays an important role, for instance to recommend music for daily activities [75].

An overview over aspects, techniques, and challenges of music recommender systems can be found in [58]. Therefore, in this section, we do not elaborate on the basics of music recommender systems but highlight again interfaces that focus on personalization and user-centric aspects, as we consider these to be the bridge to future intelligent music listening interfaces.

### 4.1 Recommender interfaces

Although most related studies have focused on methods and algorithms of music recommendation and playlist generation, some studies focus on interfaces.

*MusicSun* [50] is a user interface for artist recommendation. A user first puts favorite artist names into a "sun" metaphor, a circle in the center of the screen, and then obtains a ranked list of recommended artists. The sun is visualized with some surrounding "rays" that are labeled with words to summarize the query artists in the sun. By interactively selecting a ray, the user can look at and listen to the corresponding recommended artists.

*Moodplay* [1] is an interactive music recommender system that uses a hybrid recommendation algorithm based on mood metadata and audio content. A user first constructs a profile by entering favorite artist names and then obtains a ranked list of recommended artists. It highlights those artist positions in a latent mood space visualization, showing various mood labels. The centroid of profile artist positions is used to recommend nearby artists. The change of a user's preference is interactively modeled by moving

in this space and its trail is used to recommend artists.

In *MoodSwings* [34], users try to match each other while tracing the trajectory of music through a 2D emotion space. The users' input provides metadata on the emotional impression of songs as it changes over time.

Recent studies deal with the design of recommender user interfaces regarding complexity and user control [28] and the implications of recommender techniques on the discovery of music in playlist building [32].

## 4.2 Psychologically-inspired music recommendation

Recently, music recommender research is experiencing a boost on topics related to psychology-informed recommendation. In particular the psychological concepts of personality and affect (mood and emotion) are increasingly integrated into prototypes. The motivation for this is that while listening to music both personality traits and affective states have been shown to strongly influence music preferences [14, 52, 56].

Lu and Tintarev [41] propose a system that re-ranks results of a collaborative filtering approach according to the degree of diversity each song contributes to the recommendation list. Since previous studies showed that personality is most strongly correlated with music key, genre, and number of artists, the authors implement diversity through these features and adjust results depending on the listener's personality. Fernández-Tobías et al. [13] propose a personality-aware matrix factorization approach that integrates a latent user factor describing users' personality in terms of the Big Five/OCEAN model (openness, conscientiousness, extraversion, agreeableness, and neuroticism) [29]. Deng et al. [9] propose an emotion-aware recommender for which they extract music listening information and emotions from posts in Sina Weibo,[5] a popular Chinese microblogging service, adopting a lexicon-based approach (Chinese dictionaries and emoticons). *FocusMusicRecommender* [79] recommends and plays back musical pieces suitable to the user's current concentration level estimated from the user's behavior history.

## 5. THE NEXT PHASE: THE FUTURE OF INTELLIGENT MUSIC USER INTERFACES

In terms of interfaces, we observe strong trends towards context-awareness and personalization, also on the level of individual user and personality traits that should guide the recommendation process when other sufficient interaction data is unavailable. The central challenge behind these facets is to accurately infer the user's intent in an action (listening, skipping, etc.), i.e., to uncover the reasons why humans indulge in music, from the comparatively limited signal that is received.

On the other hand, we see the development in the realm of music generation and variation algorithms, which permit to create content based on large repositories of examples (cf. recent work by *Google Magenta*[6] [26, 53, 54])

and/or with the help of informed rules and templates, e.g., for automatic video soundtrack creation or adaptive music generation in video games marketed by companies such as *Jukedeck*,[7] or *Melodrive*,[8] respectively.

In the long run, we expect the border of these domains to blur, i.e., there will be no difference in accessing existing, recorded music and music automatically created by the system tailored to the listener's needs. More concretely, as discussed as one of grand challenges in MIR in [19], we envision music streaming systems that deliver preferred content based on the user's current state and situational context, automatically change existing music content to fit the context of the user, e.g., by varying instruments, arrangements, or tempo of the track,[9] and even create new music based on the given setting.

With the current knowledge of streaming platforms about a user's preferences, context sensing devices running the music apps, and first algorithms to variate and generate content, the necessary ingredients for such a development seem to be available already.

## 6. CONCLUSIONS

We identified three phases of listening culture and discussed corresponding intelligent interfaces. Interfaces pertaining to the first phase focus on structuring and visualizing smaller scale music collections, such as personal collections or early digital sales repositories. In terms of research prototypes, this phase is most driven by content-based MIR algorithms. The second phase deals with web-based interfaces and information systems, with a strong focus on textual descriptions in the form of collaborative tags. MIR research during this phase therefore deals with automatic tagging of music and utilization of tag information in interfaces. Finally, the third and current phase is shaped by lean back experiences driven by automatic playlist algorithms and personalized recommendation systems. MIR research is therefore shifting towards exploitation of user interaction data, however always with a focus on integration of content-based methods, community metadata, user information, and contextual information of the user. While the former three strategies are typically applied to remedy cold start problems, integrating context-awareness is often an additional source thereof.

The trend of personalizing listening experiences leads us to belief that, in the not too distant future, music listening will not only be a matter of delivering the right music at the right time, but also of generating and "shaping" the right music for the situation the user is in. We will therefore see a confluence of music retrieval and (interactive) music generation – with ample challenges for MIR research ahead.

---

[5] http://weibo.com
[6] https://magenta.tensorflow.org

[7] https://www.jukedeck.com
[8] https://melodrive.com
[9] One of the earliest approaches to customize or personalize existing music is *"music touch-up"* [18], where several examples such as *Drumix* [80] and *AutoMashUpper* [8] were developed. Lamere's *Infinite Jukebox* (http://infinitejukebox.playlistmachinery.com) can also be seen as an example toward this direction.

## 7. REFERENCES

[1] Ivana Andjelkovic, Denis Parra, and John O'Donovan. Moodplay: Interactive mood-based music discovery and recommendation. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization (UMAP 2016)*, pages 275–279, 2016.

[2] Luke Barrington, Damien O'Malley, Douglas Turnbull, and Gert Lanckriet. User-centered design of a social game to tag music. In *Proceedings of the ACM SIGKDD Workshop on Human Computation (HCOMP 2009)*, pages 7–10, 2009.

[3] Thierry Bertin-Mahieux, Douglas Eck, François Maillet, and Paul Lamere. Autotagger: A model for predicting social tags from acoustic features on large music databases. *Journal of New Music Research*, 37(2):115–135, 2008.

[4] Geoffray Bonnin and Dietmar Jannach. Automated generation of music playlists: Survey and experiments. *ACM Computing Surveys*, 47(2):26:1–26:35, 2014.

[5] Donald Byrd and Michael Fingerhut. The history of ISMIR – a short happy tale. *D-Lib Magazine*, 8(11), 2002.

[6] Michael Casey, Remco Veltkamp, Masataka Goto, Marc Leman, Christophe Rhodes, and Malcolm Slaney. Content-based music information retrieval: Current directions and future challenges. *Proceedings of the IEEE*, 96(4):668–696, 2008.

[7] Òscar Celma. *Music Recommendation and Discovery – The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer, 2010.

[8] Matthew E. P. Davies, Philippe Hamel, Kazuyoshi Yoshii, and Masataka Goto. AutoMashUpper: Automatic creation of multi-song music mashups. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(12):1726–1737, 2014.

[9] Shuiguang Deng, Dongjing Wang, Xitong Li, and Guandong Xu. Exploring user emotion in microblogs for music recommendation. *Expert Systems with Applications*, 42(23):9284–9293, 2015.

[10] Michael Dittenbach, Dieter Merkl, and Andreas Rauber. Hierarchical clustering of document archives with the growing hierarchical self-organizing map. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN 2001)*, 2001.

[11] J. Stephen Downie. The scientific evaluation of music information retrieval systems: Foundations and future. *Computer Music Journal*, 28:12–23, 2004.

[12] Douglas Eck, Paul Lamere, Thierry Bertin-Mahieux, and Stephen Green. Automatic generation of social tags for music recommendation. In *Advances in Neural Information Processing Systems 20 (NIPS 2007)*, 2008.

[13] Ignacio Fernández-Tobías, Matthias Braunhofer, Mehdi Elahi, Francesco Ricci, and Iván Cantador. Alleviating the new user problem in collaborative filtering by exploiting personality information. *User Modeling and User-Adapted Interaction*, 26(2-3):221–255, 2016.

[14] Bruce Ferwerda, Marko Tkalcic, and Markus Schedl. Personality traits and music genres: What do people prefer to listen to? In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP 2017)*, pages 285–288, 2017.

[15] Arthur Flexer, Dominik Schnitzer, Martin Gasser, and Gerhard Widmer. Playlist generation using start and end songs. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, pages 173–178, 2008.

[16] Hiromasa Fujihara, Masataka Goto, Tetsuro Kitahara, and Hiroshi G. Okuno. A modeling of singing voice robust to accompaniment sounds and its application to singer identification and vocal-timbre-similarity-based music information retrieval. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(3):638–648, 2010.

[17] Martin Gasser and Arthur Flexer. FM4 Soundpark: Audio-based music recommendation in everyday use. In *Proceedings of the 6th Sound and Music Computing Conference (SMC 2009)*, 2009.

[18] Masataka Goto. Active music listening interfaces based on signal processing. In *Proceedings of the 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2007)*, pages 1441–1444, 2007.

[19] Masataka Goto. Grand challenges in music information research. In Meinard Muller, Masataka Goto, and Markus Schedl, editors, *Dagstuhl Follow-Ups: Multimodal Music Processing*, pages 217–225. Dagstuhl Publishing, 2012.

[20] Masataka Goto and Roger B. Dannenberg. Music interfaces based on automatic music signal analysis: New ways to create and listen to music. *IEEE Signal Processing Magazine*, 36(1):74–81, 2019.

[21] Masataka Goto and Takayuki Goto. Musicream: Integrated music-listening interface for active, flexible, and unexpected encounters with musical pieces. *IPSJ (Information Processing Society of Japan) Journal*, 50(12):2923–2936, 2009.

[22] Sten Govaerts and Erik Duval. A web-based approach to determine the origin of an artist. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, 2009.

[23] Masahiro Hamasaki, Masataka Goto, and Tomoyasu Nakano. Songrium: Browsing and listening environment for music content creation community. In *Proceedings of the 12th Sound and Music Computing Conference (SMC 2015)*, pages 23–30, 2015.

[24] Negar Hariri, Bamshad Mobasher, and Robin Burke. Context-aware music recommendation based on latent-topic sequential patterns. In *Proceedings of the 6th ACM Conference on Recommender Systems (RecSys 2012)*, pages 131–138, 2012.

[25] David Hauger and Markus Schedl. Exploring geospatial music listening patterns in microblog data. In *Proceedings of the 10th International Workshop on Adaptive Multimedia Retrieval (AMR 2012)*, 2012.

[26] Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Noam Shazeer, Ian Simon, Curtis Hawthorne, Andrew Dai, Matt Hoffman, Monica Dinculescu, and Douglas Eck. Music transformer: Generating music with long-term structure. In *Proceedings of the 7th International Conference on Learning Representations (ICLR 2019)*, 2019.

[27] Sebastian Huber, Markus Schedl, and Peter Knees. nepDroid: An intelligent mobile music player. In *Proceedings of the ACM International Conference on Multimedia Retrieval (ACM ICMR 2012)*, 2012.

[28] Yucheng Jin, Nava Tintarev, and Katrien Verbert. Effects of personal characteristics on music recommender systems with different levels of controllability. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys 2018)*, pages 13–21, 2018.

[29] Oliver P. John, Eileen M. Donahue, and Robert L. Kentle. The big five inventory—versions 4a and 54, 1991.

[30] Carles F. Julià and Sergi Jordà. SongExplorer: A tabletop application for exploring large collections of songs. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, 2009.

[31] Tetsuya Kageyama, Kazuhiro Mochizuki, and Yosuke Takashima. Melody retrieval with humming. In *Proceedings of the 1993 International Computer Music Conference (ICMC 1993)*, pages 349–351, 1993.

[32] Iman Kamehkhosh, Geoffray Bonnin, and Dietmar Jannach. Effects of recommendations on the playlist creation behavior of users. *User Modeling and User-Adapted Interaction*, 2019.

[33] Joon Hee Kim, Brian Tomasik, and Douglas Turnbull. Using artist similarity to propagate semantic information. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, 2009.

[34] Youngmoo E. Kim, Erik M. Schmidt, and Lloyd Emelle. MoodSwings: A collaborative game for music mood label collection. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, pages 231–236, 2008.

[35] Peter Knees and Markus Schedl. A survey of music similarity and recommendation from music context data. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 10(1), 2013.

[36] Peter Knees, Markus Schedl, Tim Pohle, and Gerhard Widmer. An innovative three-dimensional user interface for exploring music collections enriched with meta-information from the web. In *Proceedings of the 14th ACM International Conference on Multimedia (ACM Multimedia 2006)*, 2006.

[37] Paul Lamere. Social tagging and music information retrieval. *Journal of New Music Research*, 37(2):101–114, 2008.

[38] Paul Lamere and Douglas Eck. Using 3D visualizations to explore and discover music. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, 2007.

[39] Edith L. M. Law, Luis von Ahn, Roger B. Dannenberg, and Mike Crawford. TagATune: A game for music and sound annotation. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, pages 361–364, 2007.

[40] Stefan Leitich and Martin Topf. Globe of Music - music library visualization using GeoSOM. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, 2007.

[41] Feng Lu and Nava Tintarev. 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, co-located with ACM Conference on Recommender Systems (RecSys 2018)*, 2018.

[42] Dominik Lübbers and Matthias Jarke. Adaptive multimodal exploration of music collections. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, 2009.

[43] Michael I. Mandel and Daniel P.W. Ellis. A web-based game for collecting music metadata. *Journal of New Music Research*, 37(2):151–165, 2008.

[44] Michael I. Mandel, Razvan Pascanu, Douglas Eck, Yoshua Bengio, Luca M. Aiello, Rossano Schifanella, and Filippo Menczer. Contextual tag inference. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 7S(1):32:1–32:18, 2011.

[45] Brian McFee and Gert Lanckriet. The natural language of playlists. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 2011.

[46] Fabian Mörchen, Alfred Ultsch, Mario Nöcker, and Christian Stamm. Databionic visualization of music collections according to perceptual distance. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, 2005.

[47] Robert Neumayer, Michael Dittenbach, and Andreas Rauber. PlaySOM and PocketSOMPlayer, alternative interfaces to large music collections. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, 2005.

[48] Elias Pampalk, Simon Dixon, and Gerhard Widmer. Exploring music collections by browsing different views. *Computer Music Journal*, 28(2):49–62, 2004.

[49] Elias Pampalk and Masataka Goto. MusicRainbow: A new user interface to discover artists using audio-based similarity and web-based labeling. In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, 2006.

[50] Elias Pampalk and Masataka Goto. MusicSun: A new approach to artist recommendation. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, 2007.

[51] Tim Pohle, Peter Knees, Markus Schedl, Elias Pampalk, and Gerhard Widmer. "Reinventing the Wheel": A novel approach to music player interfaces. *IEEE Transactions on Multimedia*, 9(3):567–575, 2007.

[52] Peter J. Rentfrow and Samuel D. Gosling. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6):1236–1256, 2003.

[53] Adam Roberts, Jesse Engel, Sageev Oore, and Douglas Eck. Learning latent representations of music to generate interactive musical palettes. In *Proceedings of the 2018 ACM Workshop on Intelligent Music Interfaces for Listening and Creation (MILC 2018)*, 2018.

[54] Adam Roberts, Jesse Engel, Colin Raffel, Curtis Hawthorne, and Douglas Eck. A hierarchical latent vector model for learning long-term structure in music. In *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)*, pages 4364–4373, 2018.

[55] Shoto Sasaki, Kazuyoshi Yoshii, Tomoyasu Nakano, Masataka Goto, and Shigeo Morisihama. LyricsRadar: A lyrics retrieval system based on latent topics of lyrics. In *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR 2014)*, pages 585–590, 2014.

[56] Markus Schedl, Emilia Gómez, Erika Trent, Marko Tkalčič, Hamid Eghbal-Zadeh, and Agustín Martorell. On the interrelation between listener characteristics and the perception of emotions in classical orchestra music. *IEEE Transactions on Affective Computing*, 9:507–525, 2018.

[57] Markus Schedl, Christian Höglinger, and Peter Knees. Large-scale music exploration in hierarchically organized landscapes using prototypicality information. In *Proceedings of the ACM International Conference on Multimedia Retrieval (ACM ICMR 2011)*, 2011.

[58] Markus Schedl, Peter Knees, Brian McFee, Dmitry Bogdanov, and Marius Kaminskas. Music Recommender Systems. In Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, editors, *Recommender Systems Handbook*, pages 453–492. Springer, 2nd edition, 2015.

[59] Markus Schedl, Gerhard Widmer, Peter Knees, and Tim Pohle. A music information system automatically generated via web content mining techniques. *Information Processing & Management*, 47:426–439, 2011.

[60] Dominik Schnitzer, Tim Pohle, Peter Knees, and Gerhard Widmer. One-touch access to music on mobile devices. In *Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia (MUM 2007)*, pages 103–109, 2007.

[61] Malcolm Slaney. Web-scale multimedia analysis: Does content matter? *IEEE MultiMedia*, 18(2):12–15, 2011.

[62] Malcolm Slaney and William White. Similarity based on rating data. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, pages 479–484, 2007.

[63] Mohamed Sordo. *Semantic Annotation of Music Collections: A Computational Approach*. PhD thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2012.

[64] Rebecca Stewart, Mark Levy, and Mark Sandler. 3D interactive environment for music collection navigation. In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, 2008.

[65] Sebastian Stober and Andreas Nürnberger. MusicGalaxy — an adaptive user-interface for exploratory music retrieval. In *Proceedings of the 7th Sound and Music Computing Conference (SMC 2010)*, pages 23–30, 2010.

[66] Takumi Takahashi, Satoru Fukayama, and Masataka Goto. Instrudive: A music visualization system based on automatically recognized instrumentation. In *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR 2018)*, pages 561–568, 2018.

[67] Kosetsu Tsukuda, Keisuke Ishida, and Masataka Goto. Lyric Jumper: A lyrics-based music exploratory web service by modeling lyrics generative process. In *Proceedings of the 18th International Society for Music Information Retrieval Conference (ISMIR 2017)*, pages 544–551, 2017.

[68] Douglas Turnbull, Luke Barrington, David Torres, and Gert Lanckriet. Towards musical query-by-semantic-description using the CAL500 data set. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (ACM SIGIR 2007)*, 2007.

[69] George Tzanetakis, Georg Essl, and Perry Cook. Automatic musical genre classification of audio signals. In *Proceedings of the 2nd International Symposium on Music Information Retrieval (ISMIR 2001)*, pages 205–210, 2001.

[70] Alfred Ultsch and Hans Peter Siemon. Kohonen's self-organizing feature maps for exploratory data analysis. In *Proceedings of the International Neural Network Conference (INNC 1990)*, pages 305–308, 1990.

[71] Beatrix Vad, Daniel Boland, John Williamson, Roderick Murray-Smith, and Peter Berg Steffensen. Design and evaluation of a probabilistic music projection interface. In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR 2015)*, pages 134–140, 2015.

[72] Laurens van der Maaten and Geoffrey Hinton. Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

[73] Rob van Gulik and Fabio Vignoli. Visual playlist generation on the artist map. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, pages 520–523, 2005.

[74] Shankar Vembu and Stephan Baumann. A self-organizing map based knowledge discovery for music recommendation systems. In *Proceedings of the 2nd International Symposium on Computer Music Modeling and Retrieval (CMMR 2004)*, 2004.

[75] Xinxi Wang, David Rosenblum, and Ye Wang. Context-aware mobile music recommendation for daily activities. In *Proceedings of the 20th ACM International Conference on Multimedia (ACM Multimedia 2012)*, pages 99–108, 2012.

[76] Brian Whitman and Daniel P. W. Ellis. Automatic record reviews. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004)*, pages 470–477, 2004.

[77] Erling Wold, Thom Blum, Douglas Keislar, and James Wheaton. Content-based classification, search, and retrieval of audio. *IEEE MultiMedia*, 3(3):27–36, 1996.

[78] Yingxin Wu and Masahiro Takatsuka. Spherical self-organizing map using efficient indexed geodesic data structure. *Neural Networks*, 19(6–7):900–910, 2006.

[79] Hiromu Yakura, Tomoyasu Nakano, and Masataka Goto. FocusMusicRecommender: A system for recommending music to listen to while working. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces (ACM IUI 2018)*, pages 7–17, 2018.

[80] Kazuyoshi Yoshii, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno. Drumix: An audio player with functions of realtime drum-part rearrangement for active music listening. *IPSJ (Information Processing Society of Japan) Journal*, 48(3):1229–1239, 2007.