(De)Composing the Algorithm: Explaining Music Recommender Systems to Artists for Understanding, Transparency, and Empowerment

Zuzanna Michalewicz, Karlijn Dinnissen, Eelco Herder and Hanna Hauptmann

Utrecht University, Department of Information and Computing Sciences, Utrecht, The Netherlands

Abstract

Music recommender systems (MRS) play a central role in shaping how audiences discover music. While there is growing interest in explainable and user-centered recommendations, the perspective of artists — as creators affected by these systems — remains underexplored. This study focuses on artists' information needs, their understanding of the mechanisms behind MRS, and whether providing information about how MRS work can influence the artists' perception of transparency as well as their engagement. To address these questions, we conducted a mixed-method study, which included semi-structured interviews, co-design sessions, and a questionnaire with artists. The findings suggest that while many artists have a general understanding of music recommendations from a listener's perspective, they often struggle to apply that knowledge as creators. Explanations were found to support understanding and encourage reflection on the systems' influence on their visibility and reach. Based on these insights, a prototype was developed and evaluated with a subset of the participants. Our results highlight the importance of artists understanding MRS. Additionally, the results indicate that participatory design may serve as a source of empowerment for artists.

Keywords

Music recommender systems, Transparency, Participatory design, User study

1. Introduction

Over the past two decades, recommender systems have transformed the digital landscape, changing how users discover and consume content across multiple sectors, including entertainment [1]. Within the music industry in particular, these solutions influence the content recommended to users, as streaming platforms like Spotify, Apple Music, and Tidal have gained popularity [2]. The introduction of recommender systems has changed the landscape of music discovery for listeners, providing personalized recommendations from vast catalogs [2]. Despite the fact that those systems have been widely adopted and seem successful in enhancing user experience [3], music recommender systems (MRS) research often overlooks the perspective of artists, whose livelihoods depend on these algorithmic visibility mechanisms [4]. Recent research conducted in MRS has highlighted artists' calls for greater transparency in how these systems operate, with study outcomes suggesting that artists perceive current platforms as lacking in terms of fairness and sufficient insight into recommendation mechanisms [5, 6]. This lack of transparency is particularly problematic, given that artists' professional success grows more dependent on algorithmic decisions they cannot understand or influence. To the best of our knowledge, there have been no previously published studies exploring the topic of explanations of MRS for artists.

As a first step towards improving MRS transparency for artists, this study explores four research questions:

- **RQ1**: What are artists' information needs and how would the artists like them to be addressed?

- **RQ2**: How does providing music artists with explanations on MRS affect their sense of understanding of such systems?
- **RQ3**: How does participating in co-design sessions influence the artists' sense of empowerment if they were given an opportunity to interact with music streaming platform creators in the future?
- **RQ4**: How do artists evaluate a prototype incorporating research findings from RQ1, RQ2, and RQ3 in terms of perceived usefulness and technology acceptance?

The main contributions are as follows: i) **Artist perspective insights**: empirical evidence of artists' knowledge gaps, information needs, and relationship with MRS, providing the first systematic understanding of how creators experience algorithmic music platforms, ii) **Educational framework**: development and validation of interactive educational materials that successfully improve artists' understanding of MRS mechanisms, iii) **Participatory methodology**: demonstration that co-design sessions enhance artists' sense of empowerment and ability to communicate about MRS development, iv) **Design insights**: identification of artists' prioritized features (discovery routes, engagement metrics) and creation of a prototype dashboard achieving high user acceptance scores, and v) **Methodological contribution**: a replicable three-phase approach combining interviews, education, and co-design for studying creator perspectives on algorithmic systems.

2. Literature Review

This section reviews the foundational concepts underlying this research, including MRS, explainable AI (XAI), and participatory design methodologies.

2.1. Music Recommender Systems

One of the main objectives of recommender systems is to generate meaningful recommendations for its end users [7]. Herein, different types of content demand unique approaches to recommendations. Music differs from other forms of media, such as movies, in that items are relatively short and can be consumed repeatedly [8]. Additionally, music is often consumed in sequential and curated order, such as albums or playlists. This requires a unique recommendation strategy.

When evaluating MRS and their intended purpose and functions, it is crucial to recognize the different stakeholder groups [9]. They often have different needs and goals, which can influence the objectives applicable to other parties involved. In the sphere of MRS, there are three distinct types of stakeholders: users, item providers and the streaming platforms.

The *users*, also referred to as consumers or customers, are the group that consumes the content on the platform. This group primarily seeks convenient access to the vast library, with recommendations that match their taste while suggesting new music as well. It should be noted that research on MRS has been mainly focused on this group [9]. The second group of stakeholders are the *item providers* [9], who provide the music content that is ultimately recommended (or not) to the end users. The item providers are most often artists, but also record and publishing companies [9]. Artists are considered a particularly important group among the item providers in MRS. Unlike other stakeholders, artists have a creative, and often personal, stake in how their music is discovered and recommended. Research indicates that stakeholders belonging to this group often face challenges, as their visibility directly impacts their livelihood and/or artistic recognition [9]. Their primary interests include fair representation in RS, adequate compensation for streams, and the ability to reach both existing and potential new audiences. Additionally, the artists have been found to desire greater control over the recommendation processes. In particular, the artists would like to have influence over which of their music is recommended to whom [6]. The third group of stakeholders are the *platforms* themselves [10], including Spotify, Apple Music, Amazon Music, Pandora and Tidal. These platform are considered a stakeholder as they have created proprietary MRS in order to match users with item providers, with the aim of gaining benefit [3]. The different goals of these stakeholders influence their needs in terms of transparency and explainability.

2.2. Transparency & Explainability

The term 'transparency' is often used to describe the visibility and volume of information that a service provider makes available to a particular user group [11]. There are several implementation challenges in providing transparency. Transparency inherently assumes providing other parties access to information [11]. In case of MRS, the platforms themselves would need to provide access to other stakeholders. Specifically within commercial entertainment platforms, it can be particularly challenging to obtain transparency, as the algorithms are a part of trademark secrets and are therefore kept hidden from the public eye. That can leave the artists in the dark when it comes to the functioning of the algorithm and understanding why it favors some content over other [12]. Intentional use of explainable AI (XAI) may address this information need [13] by explaining the internal system mechanics in human terms [14], therewith increasing transparency and trust [15].

Explainability is a crucial factor contributing to a system's transparency. Research on the topic suggests that the first purpose of explanations may be to gain or exchange understanding [16]. Explanations can support transparency through enabling the users to see aspects of the inner state or functionality of the AI system. Additionally, explanations can serve as decision support, helping users improve their decision-making when interacting with recommender systems [16].

2.2.1. XAI in the Music Domain

Prior work on explainable music recommendation has primarily focused on end-user experiences, exploring various explanation types for listeners, including feature-based explanations that highlight musical characteristics, collaborative explanations showing similar user preferences, and hybrid approaches combining multiple explanation modalities [12]. Research has demonstrated that users find explanations valuable for understanding why certain songs have been recommended, which can increase satisfaction and engagement with music platforms [17]. However, this body of work has predominantly addressed the information needs of music consumers rather than creators, leaving a significant gap in understanding how explanations might serve artists' distinct professional requirements.

In XAI research, often two different types of users are distinguished: experts and non-experts. XAI explanations are often created to help *machine learning experts or developers* in interpreting complex algorithms [14]. In contrast, *non-experts* (i.e., end-users) lack knowledge and understanding of how algorithmic systems works, even though they may still be knowledgeable on the domain, for instance the music genres and artists that they like or dislike.

A further distinction is that global and local explanations serve different roles in XAI for MRS [12]. Global explanations provide an overview of how a recommender system operates across the users and items, detailing how the model makes the decisions. For example, global explanations may show the principles based on which the user clusters are formed or how the system tends to recommend popular tracks to a particular demographic. On the other hand, local explanations are specific to pairs of individual items, explaining why a particular track is suggested to a particular user [12]. These explanations focus on the features or user behaviors that led to the specific recommendation, such as previous listening habits or preferences for certain genres [12, 18]. For end-users, local explanations are the most useful type of explanations.

By contrast, for music artists, global explanations could be especially beneficial because they offer insights into the overarching mechanics of the system, helping them understand not just how a single song gets recommended but how their catalog as a whole might perform on the platform. This fosters understanding of the system and increases the perceived transparency of the system to artists [12], as will be discussed in the next subsection.

2.3. Transparency for Music Artists

Some works address the needs of music artists directly, and show that artists lack understanding of how MRS work, which hinders them from effectively engaging with systems to optimize them for their

interests [6, 5].This lack of understanding further deepens the issues that artists are facing, especially when it comes to understanding how they can influence how the system deals with their music content.

Therefore, it is vital to find out what types of explanation the artists would find helpful, so that they would be able to improve their outreach on the platform and thus giving them a sense of empowerment.

As very limited previous transparency or XAI work has focused on item providers as an audience [6, 12], it also remains unclear how the information needs of this stakeholder group may be addressed through explanations, and how they should be implemented in RS. Dinnissen and Bauer [6] highlighted the need for transparency for item providers as well as agency in platform design, which directly informed the research methods utilized in this study.

3. Method

To address our research questions, we conducted a user study focusing on music artists. We employed a three-phase methodology, combining semi-structured interviews, interactive educational materials about MRS, and participatory co-design sessions to understand the artists' information needs and to develop a prototype dashboard that enhances transparency and empowerment (see Figure 1). The Ethics and Privacy Quick Scan by Utrecht University Research Institute of Information and Computing Sciences categorized the research as low-risk, meaning no additional assessment was necessary. In this section, we will describe the study setup.
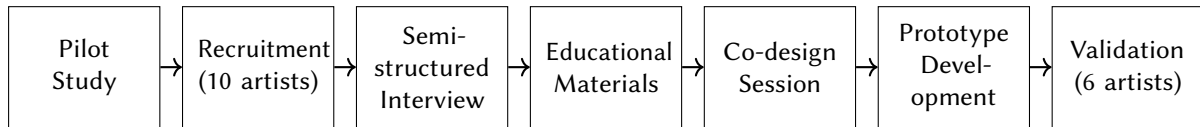

Figure 1: Study methodology overview.

3.1. Participant Recruitment and Demographics

Participants were artists, recruited through personal connections, employing a convenience sampling approach. The participants were either second or third degree connections. An overview of our participants' ages, self-identified gender, and characteristics as an artist can be found in Table 1.[1]

ID	Age	Gender	Genre	Releases	Exp.	Label	Status
P1	25-35	M	Indiepop	Singles/EP	6m-1y	No	Newcomer
P2	18-24	F	Punk, Emo	1 album	<6m	Yes	Newcomer
P3	25-35	M	Intimate funk	1 album	3-4y	Past	Newcomer
P4	25-35	M	Instrumental funk	5+ albums	5+y	Yes	Intl.
P5	25-35	M	Pop	Singles/EP	3-4y	No	Newcomer
P6	18-24	F	Indie pop	Songs/EP	6m-1y	No	Newcomer
P7	56-65	M	Rockabilly	5+ albums	5+y	Yes	Intl.
P8	36-45	M	Alt rock	3 albums	5+y	No	Intl.
P9	56-65	F	Pop/rock	2 albums	5+y	No	Other
P10	18-24	M	House	Singles/EP	1-2y	No	Newcomer

Table 1
Demographic and Professional Information of Study Participants

3.2. Knowledge Assessment Survey

After having filled out an informed consent form, participants were asked to respond to three statements on a 5-point Likert scale, varying from 1 (strongly disagree) to 5 (strongly agree). The following questions

[1]Status categories adapted from Dinnissen and Bauer [6]. 'Newcomer' indicates locally known artists with limited reach.

were asked to briefly assess how the participants evaluate their level of knowledge on MRS, whether the knowledge they have translates to understanding how MRS work, and whether understanding enables participants to apply that knowledge to their own specific case. The three statements were:

1. "I am knowledgeable about Music Recommender Systems."
2. "I understand how Music Recommender Systems work."
3. "I understand how my music is being recommended to listeners on streaming platforms."

3.3. Interview Setup and Analysis

A semi-structured interview was employed to allow for flexibility while ensuring consistent coverage of key topics across all participants. The research design was partially informed by the work of Dinnissen and Bauer [6], who previously examined item provider perspectives on influence and fairness of music streaming platforms. Specifically, questions assessing artists' understanding of recommendation models were directly adapted from their study. Table 2 shows an overview of all questions.

Q#	Question
Q1	What are your experiences with music streaming platforms as an artist?
Q2	How do music streaming platforms currently influence your career, and what was or would have been different without them?
Q3	How would you describe your level of knowledge on Music Recommender Systems
Q4	What would you say is your understanding of Music Recommender Systems?
Q5	To what extent do you think the MRS influences the way music is being consumed?
Q6	What types of information provided in your artist profile on music streaming services do you use most often?
Q7	Is there any particular information on MRS you'd be interested to know more about?
Q8	What do you think would improve your understanding of MRS?
Q9	How do you think having information on MRS would influence how you proceed with uploading and promoting your music?

Table 2
Semi-structured Interview Questions

The interview transcript analysis was conducted with the use of NVivo software. Organizing the analysis around the questions provided a deductive foundation for identifying recurring patterns. Building upon this framework, an iterative inductive process of thematic analysis was executed [19].

3.4. Educational Materials & Quiz

During the second part of the study, participants were provided access to educational materials on MRS. The materials were implemented as an interactive prototype, designed as post-hoc global explanations targeting non-expert users (artists with limited technical AI knowledge). Following literature suggesting that explanations must extend users' prior knowledge [20], the materials connected MRS concepts to the artists' professional contexts. Participants were provided with a link to the prototype and asked to share their screens and think aloud. They were informed that this part of the study would likely take 10 to 20 minutes, but we allowed them to progress at their own pace. All information was displayed textually due to feedback in the pilot study about difficulties to follow an oral presentation. The materials consisted of three parts, based on artist needs identified during the pilot study.

1. 'Explaining Music Recommender Systems'
2. 'How Interactions Shape Recommendations'
3. 'Artist Similarity: Mapping Your Musical Neighborhood'

'Explaining Music Recommender Systems'. The first section focused on explaining what MRS in general are, and gave an overview of the fundamental aspects. This included the definition of MRS as

digital matchmakers between songs and listeners, and detailed descriptions of the three primary algorithmic approaches: Content-Based Filtering (analyzing audio characteristics), Collaborative Filtering (utilizing listener networks), and Hybrid Systems (combining both methods). This section was concluded by a self-guided exploration of an 'Audio DNA' page, which offered the participants explanations of sound features used to analyze musical content, such as danceability, energy, valence, and tempo. Each feature was displayed on a card that included the feature name, the scale used to assess it, and an explanation as to what each of the features refers to. The participants could click on each card and access a list of 10 well-known songs—5 with a high value and 5 with a low value—chosen to illustrate each audio characteristic.

'How Interactions Shape Recommendations'. The second section of the prototype was designed to help the artists understand the behavioral components of MRS, focusing on how user interactions shape recommendations. It presented the MRS as 'pattern detectives' that study user engagement patterns rather than song characteristics. The interface explains how different interactions convey specific meanings: skips suggest rejection, saves indicate strong preference, repeated plays demonstrate connection, and playlist additions show contextual appreciation. The prototype breaks down how patterns are identified through three categories: immediate signals (track completion, replays, listening time), contextual clues (time of day, playlist context, device type), and continuous learning processes that refine understanding of user preferences over time.

An interactive 'Song Rating Demo', see Figure 2, allowed the users to engage with example songs from known artists to understand how collaborative filtering operates in practice. In Step 1, the participant's were asked to imagine that they are the user and simply vote for each song using thumbs up/down buttons. The system required them to address 3 out of 5 songs, in case they weren't familiar with some of them. Once the participants voted for at least 3 songs, they were shown Step 2, titled 'The Algorithm Finds Similar Users'. In this step, the artists could choose one of two profiles that had an assigned similarity score. Once the participants clicked on a chosen profile, they were shown two songs that this user might like, based on the similarities between their profiles.
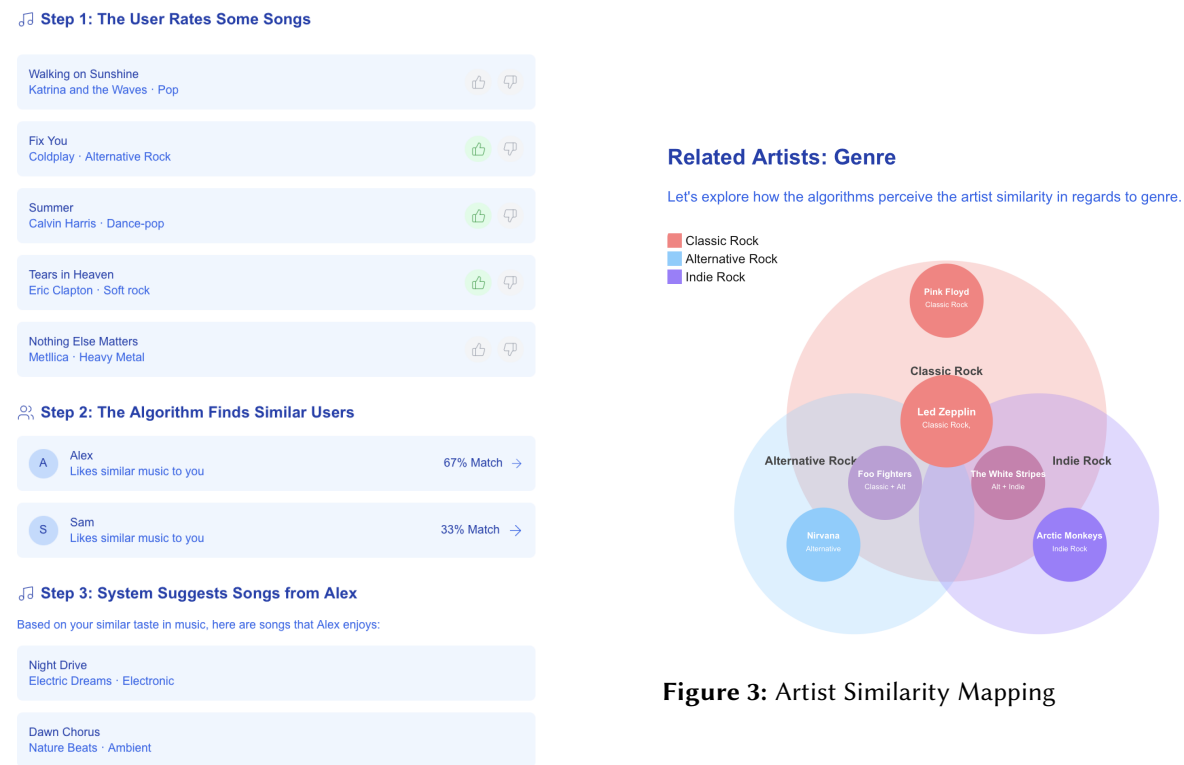
Figure 3: Artist Similarity Mapping

Figure 2: Song Rating Demo Screen

'Artist Similarity: Mapping Your Musical Neighborhood'. The final section focused on artist

similarity, explaining how MRS categorize artists and position them within the a general musical landscape. The section focused on two visualizations of artist similarity. At first, the participants were shown a Venn diagram that explained the inter-genre relationships across classic, alternative, and indie rock (Figure 3). Within each category, multiple artists were showcased, including some artists that were classified as 'in between' genres. The second displays a network diagram centered on Arctic Monkeys, connecting them to artists with varying popularity scores on Spotify's 1-100 scale [21]. The popularity is represented by the circle size and color. Together, these visualizations gave the participants an idea of how artists are classified across multiple areas.

Educational materials: think-aloud analysis. Every session was recorded from the semi-structured interviews till the end, including the part where the participants interacted with educational materials. The transcripts were later analyzed to identify where participants mentioned or asked about specific elements during the educational section. This created additional insights of where the participants might have struggled or needed clarifications.

Quiz setup. Following the presentation of the first two educational sections on MRS, participants completed a 5-question quiz designed to assess their comprehension of the key concepts covered (see Table 3). The quiz questions were derived from the contents of educational materials. While they did not check whether the participants were previously acquainted with the concepts, they were there to assess whether participants remembered the content they had read and to ensure that all participants had a basic level of understanding going into the co-design sessions.

3.5. Card Sorting & Co-Design Session

The final part of the study was a co-design session, which consisted of a card sorting exercise and a design task. This phase was designed to actively engage participants in expressing their own information needs as artists, while considering the insights they had gained from the educational materials provided earlier in the study. They were then asked to brainstorm how they would like to see their information needs integrated into a user interface in practice.

The co-design sessions were facilitated by Figjam, an interactive collaborative platform selected for its simplicity of use. Participants engaged in two tasks designed to get further insights about their preferences regarding the type of information they would like to get when it comes to MRS.

The first task (5-10 minutes) employed card sorting, where participants were asked to select and rank 10 cards from any of the 5 categories showing different types of information used in music recommendations. The five categories were: 'Listener Interaction Data', 'Artist Categorization', 'Discovery Metrics', 'Sound Profile Features', and 'Recommendation Context'. Participants assigned unique rankings (1-10) to their selected cards, based on criteria including audience understanding, music discovery potential, artist growth value, and personal goals.

In the second task (15-20 minutes), participants were asked to focus on their top 3-5 ranked cards from the first exercise and to conceptualize how they would like to see this information presented. This involved sketching and describing visualization types (charts, graphs, maps), desired level of detail, update frequency, and interaction models. Participants were encouraged to consider actionable features without concern for technical feasibility, using drawings and diagrams to communicate their ideas. The instructions asked the participants to think critically about their own circumstances, challenges, and goals, prompting them to identify specific needs that weren't being addressed by existing tools.

4. Results

Here, we present the results from the study steps mentioned in Section 3.

4.1. Knowledge Assessment Survey

The survey revealed insights on participants' self-assessed knowledge of MRS. According to the results collected, 60% of the participants initially classified themselves as 'somewhat knowledgeable' on MRS

technology. The remaining 40% either acknowledged having little to no familiarity with MRS systems or neither agreed nor disagreed that they have a level of expertise on the subject of MRS. Notably, when asked about it after completing participation in the study, 9 out of 10 participants have somewhat agreed that they are knowledgeable on MRS, bringing the baseline up significantly.

Participants were also asked about their understanding of how MRS work. Half of them expressed some degree of confidence in their understanding ('somewhat agreed'). However, the other half of participants selected 'somewhat disagree' (30%) or 'strongly disagree' (20%), signifying a lack of understanding regarding how these systems work in practice. Interestingly enough, when asked again after the study about the understanding of MRS and understanding the implications for their own music carried by MRS, the participants either strongly agreed (70%) when it came to general understanding or somewhat agreed (30%). With understanding the implications for their own music, 70% somewhat agreed while 30% strongly agreed.

This indicates a higher perception of understanding than knowledge after the study. This finding could indicate that the educational materials and explanations were successful in fostering understanding while not influencing the perceived level of knowledge.

4.2. Interview Insights

The interview analysis revealed five distinct themes, which will be described in the next subsections. These themes emerged organically from the data while maintaining clear connections to our original research questions.

General experiences with platforms. The artists' general experiences with streaming platforms reveal diverse perspectives across 90 references from all participants. The impact of streaming services on the artists' business models was shared by four different participants, with some being unhappy about loss of physical sales, while still acknowledging the streaming platforms' ability to reach wider audiences who no longer consume CDs. Concerns regarding the algorithm and content strategy were prevalent, with artists noting that their most streamed songs often are not representative of their broader musical identity, and feeling limited by discovery algorithms.

Types of platforms. Spotify stood out as the primary and preferred service, being mentioned in all 10 interviews with 13 references in total, including statements like *"My music has only been played on Spotify"* and *"Spotify is the biggest platform for me"*. Other major platforms formed a secondary tier with 12 references across 8 interviews, including Apple Music, Bandcamp, Deezer, SoundCloud, Tidal, and YouTube, suggesting that the artists maintain presence across multiple services.

Influence on career. The profound influence of streaming platforms on artistic careers emerged as a major theme, with 76 references across all interviews. Economic impact was frequently cited, with artists mentioning decreased revenues, compared to the previous revenue model which had the sales of records in its center. The participants also noted that they feel like they *"have to have a crazy big audience to make money from streaming"* and that *"smaller artists don't make a livable wage"*. Despite financial challenges, many acknowledged platforms' value for discovery and audience development, helping with visibility and reaching new listeners, though noting *"if you're not known, they don't help"*.

Level of understanding. The artists self-reported levels of understanding on how MRS operate varied within each platform, with 27 references across 5 interviews revealing both insights and misconceptions. Some participants showed basic algorithm literacy, recognizing that MRS function based on listener activity and audience response patterns. However, considerable knowledge gaps persist, with many admitting that they *"don't really understand how algorithms work"* or how their music gets recommended to listeners outside their local area. Further, artists hold specific beliefs on how algorithms work, such as the notion that *"a lot of skips mean that the song is not good"*. These beliefs potentially have an influence on their creative decisions, based on the some artists expressing that they would adapt their creative processes to the data that they see in a dashboard.

Information needs. The artists indicated various information needs when it comes to MRS. All participants expressed knowledge gaps, with a total of 51 references across interviews. Artists were particularly curious about how the MRS actually work. They wanted to know *"how the music gets*

selected", referred to the *"secret code for recommender systems"*, and asked *"which variables influence music recommendation"*. Many artists were interested in getting insights on specific factors influencing the recommendations. They asked whether music suggestions are *"based on location"*, *"purely based on genre"*, or influenced by *"related artists"*. They were also interested to get more insight into listener behavior. They mentioned that they would like to know *"how long people listen"*, or *"where in the song users are skipping"*. The most frequently mentioned need (13 references) was help on promotion strategy, including how to get on playlists, increase visibility, and decide between releasing albums versus singles.

4.3. Effects of Educational Materials

The post-interaction surveys showed mixed results. When asked about difficulties understanding MRS functionality, 5 participants explicitly stated 'No', indicating good comprehension of the materials. Multiple participants mentioned the sections on algorithmic approaches as particularly valuable. One participant specifically requested *"more in-depth explanations on taste profiles of users"*.

The results of the quiz after the first part of the educational materials (see Table 3) suggest a strong understanding of MRS concepts among the participants. All participants correctly identified what danceability means. Questions about content-based filtering and valence each achieved 90% accuracy, while collaborative filtering and instrumentalness questions showed slightly lower performance at 80% correct responses.

Question	% Correct
What is danceability in music recommendation systems?	100%
What collaborative Filtering primarily relies on	80%
What the valence of a track measures	90%
What content-Based Filtering focuses on	90%
What instrumentalness measures:	80%

Table 3
Section 1 Quiz Results: Percentage of Correct Responses

The results of the Part 2 quiz (see Table 4) suggest a good understanding of MRS interaction concepts among participants. The question about what type of interaction might indicate that a track is 'not quite right' received perfect scores (100%), showing that this concept was clearly communicated. Questions about the primary purpose of MRS and indicators that a user loves a track both achieved 90% accuracy.

Question	% Correct
What is the primary purpose of a Music Recommender System?	90%
Which of the following interactions suggests a user loves a track?	90%
What type of interaction might indicate a track is "not quite right"?	100%
What is a significant limitation of Music Recommender Systems?	70%
Which is an example of a contextual clue used by MRS?	80%

Table 4
Section 2 Quiz Results: Percentage of Correct Responses

4.4. Card Sorting & Co-Design Session

A card selection analysis was conducted to uncover patterns in how participants prioritize different types of information when it comes to music. The scale for the card-sorting exercise was 1-10, with the participants' highest-priority card receiving 10 points and their lowest-priority card receiving 1 point. This method enabled the results to be a direct reflection of the participants' original card sorting priorities.

The card sorting exercise supported the findings from the interviews to some extent. The participants rated the cards from the Discovery Metrics and Listener Interaction the highest, indicating that those pieces of information are relevant to them. New Listener Sources was clearly the most important metric to artists, scoring on average 7.33 out of 10, much higher than any other card. This aligns with the findings from the interviews as well, where a few artists inquired about *"how the music gets selected"* and how it reaches new listeners outside their existing fan base. Notably, there was no interest in investigating specific sound features (all scored 0.00) but the artists did express that they would value having a more in-depth understanding their general sound profile (3.89). They showed little interest in industry classification systems like Popularity-based Classification (0.78) and Cross-Genre Appeal (1.11), further demonstrating that artists care more about their relationship with listeners than about how they look in comparison to 'similar' artists. A sorted overview of the average rankings is shown in Table 5.

Metric	Value
New Listener Sources	7.33
Listener Save/Like Actions	5.56
Repeat Listen Patterns	4.33
Playlist Discovery Routes	4.33
Sound Characteristics (in general)	3.89
Playlist Additions	3.89
Fan Base Overlap	3.78
Similarity to other Artists	3.67
Listener Mood Matching	3.67
Genre Classification	3.67
When Your Music is Recommended	3.44
Song Completion Rates	3.22
Skip Rates of Your Songs	2.22
Listener Demographics	2.22
Geographic Reach	2.11
Time-of-Day Recommendations	1.89
Cross-Genre Appeal	1.11
Popularity-based classification	0.78
Any specific sound characteristic	0.00

Table 5
Card Sorting Results: Individual Metric Rankings

General insights. The co-design session insights varied greatly between participants. While they were encouraged to draw how they would like to see the features chosen during card-sorting to be implemented, a few of them also left text output. Two participants expressed that they do not particularly care how the insights will be visually represented, as long as they are there. Others created sketches showing the possible visual implementation of features in the form of charts, graphs and tables. Every participant approached the task differently, which allowed for getting various design ideas for multiple features. Additionally, this part of the study served as an opportunity for the artists to creatively express what they would wish to see, iterating on the already existing classification. Beyond the pre-selected features, we collected other insights that the participants' provided, analyzed them and implemented into the final design. Following this approach enabled us to create a solid framework for the prototype based on quantitative data as well as allowed us the flexibility that comes with qualitative data.

5. Dashboard Development and Evaluation

After collecting the information needs from artists, we implemented these needs into a prototype designed to help the artists understand how their music is categorized and recommended within the MRS. Consequently, we validated this prototype with participants from the original study. The main

objective of this validation study was to evaluate the usefulness and relevance of the proposed music analytics dashboard, to assess whether the various data visualizations met the artists' information needs, and to determine if integrating these features would enhance their existing analytics tools. The dashboard format was chosen as artists were already familiar with this interface through existing streaming platform analytics tools, providing a recognizable framework for organizing information. Participants were recruited from the same group of artists who had participated in the initial research interviews, with 60% agreeing to participate in the prototype evaluation.

The evaluation was based on a survey containing both quantitative and qualitative questions. Each dashboard section was evaluated using a 5-point Likert scale addressing the satisfaction of information needs and whether they would find a specific component being integrated into a dashboard helpful.

Prior to engaging with the dashboard, participants were provided with a 10-minute guided video prototype walkthrough, demonstrating the dashboard sections and functionality.[2] Following the walkthrough, participants could explore the prototype themselves with no time limit, though self-exploration data was not captured.

5.1. Design

The dashboard prototype was developed based on the key needs identified in Section 4. There, artists consistently expressed a need for greater transparency and information that would improve their understanding of the audience that engages with their content. The design prioritizes the most highly-rated features: listener source analytics, engagement tracking, playlist discovery, fan base overlap analysis, recommendation insights, mood matching capabilities, sound characteristic analysis, and retention metrics. Those were based on the eleven top chosen cards by participants during the co-design session. Our approach aligns with participatory design methodologies that involve stakeholders directly in system development, allowing the creation of tools that better meet their needs [22].

As part of the co-design task, the participants were asked to provide sketches of their top chosen cards. We used those sketches to inform the prototype creation process, though it was not possible to implement them one on one for multiple reasons. First, they often were very high-level sketches, which did not cover interactivity or functionality aspects. Second, there were almost no repeating designs, even though there were many features that multiple participants sketched or described.

The dashboard's information architecture follows a modular, hierarchical organization, from most general information to the most detailed metrics. This order was established based on the interview outcomes, with artists mentioning the need for a general overview first, and dive into details later. This aligns with Shneiderman's visual information seeking mantra [23]. The design choice also reflects the visual explanation approach, presenting information through charts, graphs, and interactive elements rather than purely textual explanations [24].

The prototype provides post-hoc global explanations designed for non-expert users, following literature suggesting that explanations should extend the users' prior knowledge rather than focusing on technical implementations [20]. Additionally, the prototype design addresses the three key dimensions of XAI: it provides global explanations (system-wide understanding rather than individual recommendations), uses post-hoc methods (explanations added externally to existing systems), and utilizes primarily visual explanation types through interactive dashboards and data visualizations [12].

The overall structure is based on five primary modules accessible through a tab-based navigation system: Content Performance Metrics, Music Discovery Routes, Fan Base Overlap, Sound Characteristics, and Time Recommendations. See Figure 4 and Figure 5 for an exemplary screenshot of the Time Recommendations and Fan Base Overlap tabs. Each tab is explained in more detail in the next section.

Implications for streaming platform design. The outcomes of the study suggest that the prototype should prioritize transparency around discovery metrics and listener interaction data. Educational components should be integrated throughout, enabling artists to get acquainted with the concepts as they are using the dashboard. Sound characteristics should be presented in as additional insights, including a range of different ones, rather than as standalone metrics.
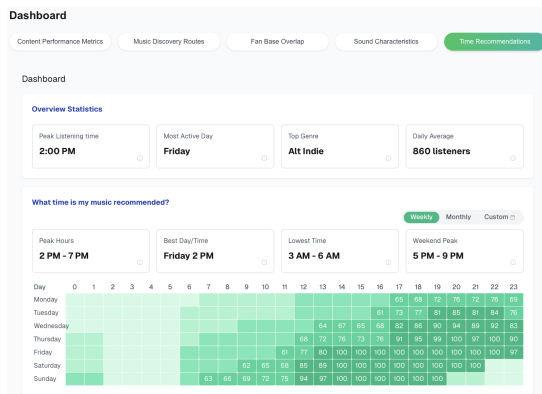
---

[2]Video available at: https://youtu.be/EUnTaCqH9pU?si=XVBQxWR6Sh9sdnjn
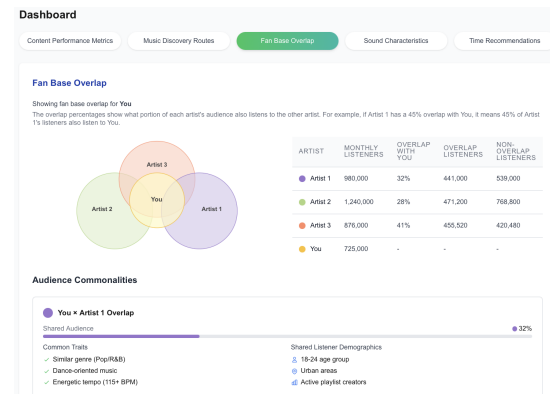
Figure 4: Time Recommendation Section

Figure 5: Fan Base Overlap Section

5.2. Key Features and Functionalities

Content performance metrics section. The Content Performance Metrics section provides an engagement information overview as well as detailed statistics for saves, listens, repeat listens and playlist additions within the time ranges of choice. Performance trends are visualized through line charts showing the performance patterns, allowing artists to understand the trends across various time frames, including daily, weekly, and monthly views. Through an interface featuring pill-shaped navigation, artists can explore different metric categories. Each of the sections included relevant information which was displayed with their respective growth trends.

Music discovery routes. The Music Discovery Routes module focuses on providing insights into how listeners find and engage with music. This section is organized into three views that progressively reveal insights into audience acquisition patterns. The New Listener Sources Overview presents a breakdown of how new listeners discover the artist's music, categorizing pathways into Recommendations, Search, Direct Links, and Other channels. The Discovery Routes section uses a stacked area chart to display how discovery patterns evolve over time, showing the proportion of listeners coming through Direct Discovery, Via Other Playlists, and Recommendations. The Playlists view provides performance data for specific playlists within which the artist's music was included.

Fan base overlap. The Fan Base Overlap module provides artists with insights into audience crossover with similar artists in the ecosystem. This section visually maps the artists' shared listeners with other bands through a Venn diagram, illustrating the intersections between the artist's audience and artists they share audience with. The table describing the diagram included metrics such as monthly listeners, the percentage of overlap between the artist and the other bands as well as the total number of the listeners that overlapped and those who have not. The Audience Commonalities section delves deeper into the shared aspects of these audiences. For each overlapping artist, the dashboard displays a bar indicating the relative strength of the connection, followed by two key categories of information: Common Traits and Shared Listener Demographics.

Sound characteristics. The Sound Characteristics module provides artists with a detailed analysis of their music, offering insights into the sound features that define their work. At the top level, six key features are displayed as numerical values in prominent metric cards: Danceability, Energy, Valence, Acousticness, Instrumentalness, and Tempo. Each metric is accompanied by an information tooltip that provides additional information about what the parameter measures to ensure full understanding of the concepts. The Song Analysis table at the bottom of the module allows artists to examine their individual tracks across all sound parameters.

Time recommendations. The Time Recommendations module offers artists strategic guidance for content release and promotion scheduling based on listener activity patterns. By informing the artists on when their music is being listened to most often, the artists can identify what periods throughout the week could potentially lead to increased visibility and tweak their release strategy towards those times. The module begins with Overview Statistics featuring four key metrics in a card format: Peak

Listening Time, Most Active Day, Top Genre, and Daily Average Listeners. The 'What time is my music recommended?' section delivers detailed temporal insights using four complementary metrics: Peak Hours, Best Day/Time, Lowest Time, and Weekend Peak. The visualization displays audience engagement patterns by hour and day of week using progressively deeper shades of green to indicate higher recommendation likelihood.

5.3. Validation Results

The prototype evaluation demonstrated strong user acceptance across all measured dimensions. The assessments of the Technology Acceptance Model revealed consistently high scores, ranging from 4.67 to 5.0 out of 5.0, with perfect scores (5.0/5.0) achieved for efficiency improvement, ease of understanding, and usability. The remaining metrics (enjoyment enhancement, learning curve, perceived value, and satisfaction) scored between 4.67-4.83, potentially indicating strong overall acceptance.

The individual dashboard sections received varying levels of validation. The Time Recommendations section achieved unanimous approval with perfect scores (5.0/5.0) for both addressing information needs and integration value, with all participants strongly agreeing on its usefulness. Music Discovery Routes also performed exceptionally well, receiving perfect scores for addressing information needs and near-perfect ratings (4.83/5.0) for integration value. The remaining sections—Content Performance Metrics and Fan Base Overlap and Sound Characteristics—all achieved positive scores above 4.0. However, the responses for these had a higher variance in scores, suggesting that they may require further refinement.

Qualitative feedback reinforced these quantitative findings, with participants expressing statements like *"I would absolutely use it"* and *"I really love the feel and interface of it all"*. When asked about future iterations, three participants indicated they would not change anything, while others requested *"even more details about the music discovery routes"* or noted that *"sound characteristics are least interesting"*, providing clear direction for future development.

Individual Section Results. The results demonstrate positive responses across all sections. An overview of the specific component scores is shown in Table 6, and the overall dashboard evaluation can be found in Table 7. The 'Time Recommendations' section was unanimously voted as the most valuable component, receiving perfect scores of 'strongly agree' for both addressing information needs and integration value from all participants. The 'Music Discovery Routes' section also received high ratings with a perfect score (5.0/5.0) for addressing information needs, and near-perfect ratings (4.83/5.0) for integration value. The average scores in the evaluation varied between 4.67 and 5.0, indicating a highly positive user perception of the prototype regarding its usefulness, ease of use, and overall acceptability.

Section	Information Needs	Integration Value
Content Performance Metrics	4.67	4.50
Music Discovery Routes	5.00	4.83
Fan Base Overlap	4.67	4.67
Sound Characteristics	4.00	4.00
Time Recommendations	5.00	5.00

Table 6
Section-wise Evaluation Results

Qualitative feedback. When asked about what should be kept the same, participants highlighted several valuable features that should be maintained in future iterations. Insights into sections like sound characteristics, fan base overlap, time recommendations and content metrics were named specifically across multiple users. Another user indicated that they would keep *"all of it"*, indicating general positive experience using the dashboard. One participant specifically appreciated the color coding, as it helps them to *"distinct the information better"*.

When asked about what should be changed, 3 participants indicated that they would not change anything. Others asked to include *"even more details about the music discovery routes"* or indicated that the *"sound characteristics are least interesting"*. One participant mentioned that the *"Content Performance*

Technology Acceptance Dimension	Average Score
Efficiency improvement	5.00
Enjoyment enhancement	4.83
Learning curve	4.67
Easy to understand	5.00
Valuable tool	4.83
User satisfaction	4.67
Easy to use	5.00

Table 7
Technology Acceptance Model Results on the overall dashboard

Metrics" tab *"already exists (almost)"*; however, this participant's opinion was not in line with the rest of the artists, who otherwise voted that they strongly agreed that they found it helpful.

When asked about any other feedback, participants expressed overall satisfaction with the prototype, praising its design and usability. Two of them responded *"I would absolutely use it"* and *"I really love the feel and interface of it all"*. One participant suggested incorporating marketing recommendations based on the metrics, highlighting a potential area for enhancement.

6. Discussion

The following section summarizes our results in the context of existing work and discusses its potential impact on future platform design and its limitations.

6.1. Key Findings

The findings reveal several insights about the artists' relationships with MRS:

Disconnect in knowledge-understanding: While 60% of participants initially perceived themselves as 'somewhat knowledgeable' about MRS, this did not translate into comprehensive understanding, with many stating they *"don't really understand how algorithms work"* or how their music gets recommended. This finding is likely applicable in other domains, such as e-commerce spaces, where the sellers might see the outcome of their listings being 'promoted' while not necessarily understanding how that happens.

Discovery-focused information priorities: Card sorting revealed that 'New Listener Sources' was the highest priority (7.33/10), with Discovery Metrics and Listener Interaction Data clusters scoring highest, while technical sound characteristics received minimal interest. These choices reflect the artists' desire for greater control, as indicated in previous work where artists expressed the wish to influence how their music is being recommended [6], as well as having a deeper understanding of their algorithmically-acquired audiences. These results indicate that the focus of future research should be put on audience relationships. The prioritization of discovery-related features could also be useful in other domains, such as reciprocal recommender systems, allowing the users to understand who is recommended to whom and why.

Evaluation of the educational material: The post-study assessments results have indicated that the educational materials have improved the understanding of MRS-related concepts. Research [15] shows that explanations can increase the understanding of AI systems, suggesting that explanations could emerge as a tool that enables the item providers to understand the recommender systems ecosystems better in other areas of the entertainment industry, like video creation.

Co-design empowerment outcomes: Participants reported feeling *"better equipped to potentially influence how MRS might recommend my music"* and appreciated the opportunity to specify desired features, indicating enhanced sense of agency through participatory processes. These results are supported by the literature [22, 25] and indicate the potential benefits of including multiple stakeholders in platform design processes, instead of focusing solely on the end users.

Prototype acceptance. Overall dashboard evaluation based on the Technology Acceptance Model

were positive, with top scores (5.0/5.0) for efficiency improvement, ease of understanding, and usability. The 'Time Recommendations' section received unanimous approval from all participants, while 'Music Discovery Routes' also performed well. The consistently high scores across both individual sections and overall dashboard metrics suggest strong user acceptance of the proposed solutions. The TAM assessments offer particularly valuable insights regarding potential adoption, with the dashboard receiving perfect scores in critical dimensions: efficiency improvement, comprehensibility, and ease of use. According to this model's principles, technology adoption depends primarily on perceived usefulness and perceived ease of use, both of which received high ratings in this evaluation.

Methodological reflections. The three-part methodology provided complimentary perspectives. The interviews established a baseline regarding knowledge. Then, the educational materials addressed knowledge gaps, ensuring that participants have the same baseline knowledge during the co-design sessions. The subsequent co-design sessions identified specific themes that the artists were interested in and allowed them to elaborate their ideas through sketching.

6.2. Implications for Platform Design

The outcomes suggest that the prototype should prioritize transparency around discovery metrics and listener interaction data, illustrating how listeners discover music through MRS and which sources generate suggestions. Additionally, educational components should be integrated throughout the platform. While platforms must balance multiple objectives, including user engagement and revenue [26], providing artist transparency could offer strategic advantages through improved content quality without requiring fundamental algorithmic changes. The hierarchical design approach mitigates information overload by presenting high-level metrics first, with modular architecture allowing artists to engage at their preferred depth and addressing varying analytical interests within the artist community.

6.3. Limitations

We acknowledge several limitations. The limited sample size limits the ability to generalize results to the broader artist population. Further, the chosen research methods could have potentially increased the selection bias and homogeneity of resulting participant pool. However, the sample achieved diversity in age (18-65), experience (less than 6 months to over 5 years), and genre representation. Additionally, recruiting participants from the initial interviews for validation could have introduced bias, as these artists were familiar with the research and had seen the educational materials. Finally, the prototype presented mock-up data rather than participants' real data, which may have influenced their perception. Allowing participants to access their own data could reveal unmet needs or demonstrate actual benefits. However, we believe that our approach is sufficient for initial prototype testing.

7. Conclusion

This work explored the information needs of artists, their understanding of system mechanisms, and the role of information on the inner workings of MRS in improving transparency and empowerment. The research aimed to understand how artists interact with these systems, what information they find meaningful, and how educational materials can support both understanding and empowerment.

The findings reveal significant gaps between the artists' perceived and actual understanding of MRS. While participants often had a general awareness from a listener's perspective, they struggled to apply this knowledge as creators. The educational interventions proved effective in improving understanding, with participants showing increased confidence in their knowledge post-study. The co-design process highlighted the value of artist participation in tool creation, resulting in a prototype that reflected their preferences and needs. Key contributions include:

1. Identification of artists' information priorities: 'New Listener Sources' emerged as the highest priority, with discovery metrics and listener interaction data ranking significantly higher than technical sound characteristics or industry classifications.

2. Educational material effectiveness indicators: The high quiz performance of artists and the positive changes of self-assessment after the study suggested that explanatory materials may contribute to improved understanding of MRS mechanisms.

3. Validation of the prototype: High Technology Acceptance Model scores (4.67-5.00) across all dimensions indicate strong potential for adoption of transparency-focused analytics tools.

4. Demonstration of participatory design value: The co-design process generated new insights by enabling artists to articulate needs that they had previously struggled to express.

These four insights suggest that MRS should embrace creator-centered design approaches, considering artist perspectives and end-user needs. As algorithms continue to shape music discovery and consumption, empowering artists becomes increasingly important. Providing them with relevant information and systems that reflect their needs deepens their understanding and opens doors for more meaningful participation in how these technologies evolve. As gatekeepers of the music industry, MRS should ensure a healthy supply side, with artists who feel empowered to continue to pursue their careers. By integrating explanatory features and opportunities for feedback, platforms could offer artists such a strengthened participation in algorithmic ecosystems.

Future work should include broader artist populations, conduct comparative analyses across multiple streaming platforms, and investigate interactive real-time explanations with the artists' own data through longitudinal studies. Additionally, research should explore marketing-oriented analytical insights based on recommendation metrics and pursue industry collaborations to test explainability interventions in live environments where all stakeholder needs are considered.

Declaration on Generative AI

During the preparation of this work, the authors employed ChatGPT for programming support and to improve the clarity of the manuscript. The generated outputs were carefully reviewed and edited, and the authors take full responsibility for the final content.

References

[1] K. Um, J. Jung, Evolution and historical review of music in mass media, International Journal of Advanced Culture Technology 12 (2024) 370–379.

[2] M. O'Dair, A. Fry, Beyond the black box in music streaming: The impact of recommendation systems upon artists, Popular Communication 17 (2019) 65–77. doi:10.1080/15405702.2019.1627548.

[3] R. Burke, Recommender systems: An overview, AI Magazine 23 (2002) 77–80.

[4] K. Dinnissen, C. Bauer, Fairness in music recommender systems: a stakeholder-centered mini review, Frontiers in Big Data 5 (2022). doi:10.3389/fdata.2022.913608.

[5] A. Ferraro, X. Serra, C. Bauer, What is fair? Exploring the artists' perspective on the fairness of music streaming platforms, 2021. doi:10.1007/978-3-030-85616-8_33.

[6] K. Dinnissen, C. Bauer, Amplifying artists' voices: Item provider perspectives on influence and fairness of music streaming platforms, in: Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization, UMAP '23, Association for Computing Machinery, New York, NY, USA, 2023, pp. 238–249. URL: https://doi.org/10.1145/3565472.3592960. doi:10.1145/3565472.3592960.

[7] P. Melville, V. Sindhwani, Recommender systems, in: C. Sammut, G. I. Webb (Eds.), Encyclopedia of Machine Learning, Springer, Boston, MA, 2010, pp. 829–838. doi:10.1007/978-0-387-30164-8_713.

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

[9] J. Kunkel, T. Donkers, User-centeredness in algorithmic music recommendations: An analysis based on interaction logs, in: Adjunct Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization (UMAP 2019), volume 2462 of *CEUR Workshop Proceedings*, 2019.

[10] M. Eriksson, R. Fleischer, A. Johansson, P. Snickars, P. Vonderau, Spotify Teardown: Inside the Black Box of Streaming Music, MIT Press, Cambridge, MA, 2019.

[11] M. Turilli, L. Floridi, The ethics of information transparency, Ethics and Information Technology 11 (2009) 105–112. doi:10.1007/s10676-009-9187-9.

[12] D. Afchar, A. Melchiorre, M. Schedl, R. Hennequin, E. Epure, M. Moussallam, Explainability in music recommender systems, AI Magazine 43 (2022) 190–208. URL: https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/21743. doi:10.1002/aaai.12056.

[13] W. Saeed, C. Omlin, Explainable ai (xai): A systematic meta-survey of current challenges and future opportunities, Knowledge-Based Systems 263 (2023) 110273.

[14] C. Sardianos, et al., The emergence of explainability of intelligent systems: Delivering explainable and personalized recommendations for energy efficiency, Intelligent Systems (2020).

[15] A. Bhattacharya, J. Ooge, G. Stiglic, K. Verbert, Directive explanations for monitoring the risk of diabetes onset: Introducing directive data-centric explanations and combinations to support what-if explorations, in: Proceedings of the 28th International Conference on Intelligent User Interfaces, IUI '23, Association for Computing Machinery, New York, NY, USA, 2023, pp. 204–219. doi:10.48550/arXiv.2302.10671.

[16] N. Tintarev, J. Masthoff, Beyond explaining single item recommendations, in: F. Ricci, L. Rokach, B. Shapira (Eds.), Recommender Systems Handbook, Springer, New York, NY, 2022. doi:10.1007/978-1-0716-2197-4_19.

[17] M. Himan, M. S. Pera, M. D. Ekstrand, Explaining recommendations in music streaming services: A comparative study of explanations as a means to promote user engagement, ACM Transactions on Interactive Intelligent Systems 13 (2023) 1–30.

[18] S. S. Y. Kim, E. A. Watkins, O. Russakovsky, R. Fong, A. Monroy-Hernández, "help me help the ai": Understanding how explainability can support human-ai interaction, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, Association for Computing Machinery, New York, NY, USA, 2023. doi:10.1145/3544548.3581001.

[19] V. Clarke, V. Braun, Thematic analysis, in: Encyclopedia of critical psychology, Springer, 2014, pp. 1947–1952.

[20] S. Coppers, J. V. den Bergh, K. Luyten, K. Coninx, I. van der Lek-Ciudin, T. Vanallemeersch, V. Vandeghinste, Intellingo: An intelligible translation environment, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), ACM, Montreal, QC, Canada, 2018, pp. 1–13. doi:10.1145/3173574.3174098.

[21] Spotify, Get artist - spotify web api reference, 2024. URL: https://developer.spotify.com/documentation/web-api/reference/get-an-artist, accessed: 2024-12-XX.

[22] T. Zamenopoulos, K. Alexiou, Co-design As Collaborative Research, Connected Communities Foundation Series, Bristol University/AHRC Connected Communities Programme, Bristol, 2018. URL: https://connected-communities.org/wp-content/uploads/2018/07/Co-Design%5fSP.pdf.

[23] B. Shneiderman, The eyes have it: A task by data type taxonomy for information visualizations, in: Visual Languages, IEEE Symposium on, IEEE Computer Society, 1996, pp. 336–336.

[24] M. Szymanski, M. Millecamp, K. Verbert, Visual, textual or hybrid: the effect of user expertise on different explanations, in: Proceedings of the 26th International Conference on Intelligent User Interfaces, IUI '21, Association for Computing Machinery, New York, NY, USA, 2021, pp. 109–119. doi:10.1145/3397481.345066.

[25] S. C. Moser, Can science on transformation transform science? lessons from co-design, Current Opinion in Environmental Sustainability 20 (2016) 106–115. doi:10.1016/j.cosust.2016.05.010, sustainability challenges.

[26] W. Bendada, G. Salha-Galvan, T. Bontempelli, Beyond accuracy: Evaluating music recommenders from a business perspective, in: Proceedings of the 17th ACM Conference on Recommender Systems, ACM, 2023, pp. 78–87.