ISSN: 3049-0952
www.jsetms.com

Comprehensive Analysis of Music Streaming Behaviour Using Spotify data and Insights

[1]Kokkirigadda Ratna Priyanka
[1]M.Tech, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
priyankaratna19@gmail.com

[2]Mr. Nandana Kumar
[2]Assistant Professor, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
nandan@dnrcet.org

Abstract*— Music streaming platforms have transformed listening from an album-oriented activity into a continuously personalized, data-driven experience. Spotify generates a rich behavioural footprint involving listening histories, track characteristics, artist preferences, playlist interactions, temporal patterns, discovery activity, and contextual signals. However, existing research frequently treats music consumption as separate problems such as recommendation, popularity prediction, user profiling, or audio-content analysis. This review synthesizes recent research on music-streaming behaviour with particular emphasis on Spotify data and identifies the limitations of isolated analytical approaches. Six interrelated dimensions are examined: multi-source music-behaviour representation; multi-temporal user preference modelling; content-driven and collaborative recommendation; contextual, social, and new-release discovery; explainability, diversity, and popularity bias; and adaptive learning for changing user preferences. Recent Spotify research demonstrates that generalized user representations can combine multimodal signals across different temporal scales and support multiple downstream personalization tasks, while studies of new-release and socially motivated listening demonstrate that recommendation relevance depends on content age, community, culture, and timing. Based on this synthesis, this review develops a Unified Spotify Behaviour Analytics Framework (USBAF) consisting of data acquisition, behavioural representation, intelligence, adaptive explanation, and insight layers. The review concludes with open challenges and a staged roadmap toward context-aware, explainable, diverse, continually adaptive, and generative music-streaming intelligence.

Keywords*— Spotify, music streaming behaviour, music recommendation, user profiling, personalization, recommender systems, music discovery, preference modelling, explainable AI, streaming analytics.

I. INTRODUCTION

The rapid adoption of on-demand music streaming has fundamentally changed how listeners discover, select, repeat, and share music. Unlike traditional music consumption, where users commonly selected albums or artists through relatively static collections, streaming platforms provide access to continuously expanding catalogues and dynamically generated recommendations. Consequently, listening behaviour has become an important source of information for understanding user preferences, predicting future consumption, improving discovery, and designing personalized recommendation systems. Spotify is a particularly relevant environment for this research because its ecosystem combines large-scale listening interactions with music metadata, audio characteristics, recommendation surfaces, playlists, search, and discovery mechanisms.

Music-streaming behaviour is nevertheless more complex than the simple relationship between a user and a preferred song. A listener may repeatedly consume familiar tracks during routine activities while simultaneously exploring newly released music during discovery sessions. Preferences may also change according to time, mood, social context, location, cultural background, or the age of the content. Recent Spotify research using listening data from one million users and 282,000 newly released tracks found that preferences for newly released music are not identical to users' overall music preferences, demonstrating that new-release recommendation should be treated as a distinct analytical problem [1].

The research literature has consequently moved from conventional content-based and collaborative-filtering approaches toward hybrid systems that combine multiple forms of information. A 2024 review of 55 studies on content-driven music recommendation identified six persistent challenges: diversity and novelty, transparency and explanations, context awareness, sequence recommendation, scalability and efficiency, and cold-start handling [2]. These challenges indicate that recommendation accuracy alone is insufficient for understanding modern music-streaming behaviour.

Another important development is the use of generalized user representations. Spotify's 2025 research describes a two-stage framework in which rich user signals are compressed into stable user embeddings and then transferred to downstream retrieval, ranking, search, and recommendation tasks. The framework aggregates information over approximately six months, one month, and one week to represent core interests, medium-term shifts, and fresh intent, respectively [3]. This development suggests a transition from independent task-specific models toward shared representations of user behaviour.

Social and cultural factors further complicate the recommendation problem. Spotify research using listening activity from 100,000 users across 20 countries examined recommendations designed around users' desire to listen to music trending within their communities. The study showed that recommendation timeliness and cultural context can influence the usefulness of socially motivated music discovery [4]. Thus, a comprehensive behavioural analysis should consider not only what a user likes, but also when, why, and within which social context the user consumes music.

This review therefore addresses the following question: How can Spotify listening data and related research be synthesized into a comprehensive framework for understanding, predicting, and adapting to music-streaming behaviour? The contributions of this review are fourfold. First, it organizes Spotify-related

ISSN: 3049-0952
www.jsetms.com

behavioural information into a unified analytical representation. Second, it synthesizes recommendation, personalization, new-release discovery, social listening, and popularity analysis as interconnected problems rather than isolated tasks. Third, it identifies explainability, diversity, popularity bias, cold-start behaviour, and preference drift as central challenges. Fourth, it proposes a five-layer Unified Spotify Behaviour Analytics Framework (USBAF) and a staged roadmap toward adaptive music-streaming intelligence.

II. MUSIC-STREAMING DATA AND BEHAVIOURAL REPRESENTATION

A. Sources of Music-Streaming Behaviour

A streaming platform provides multiple categories of information that can be used to represent listening behaviour. At the user level, historical interactions can include plays, skips, repeats, listening duration, artist consumption, playlist interaction, search behaviour, and discovery activity. At the item level, information may include artist, album, genre, release age, popularity, and audio characteristics. These sources differ substantially in their semantic meaning and temporal dynamics. Audio features describe properties of a track and can support content-driven similarity, whereas listening interactions provide collaborative evidence about how users actually consume music. Metadata supplies semantic information, while temporal information provides context about when consumption occurs. Consequently, a robust behavioural representation should avoid relying on a single data source.

Recent music-recommendation literature confirms this transition. Deldjoo et al. reviewed content-driven recommendation approaches and organized music information into layers ranging from the audio signal to metadata, expert-generated content, user-generated content, and derivative content [2]. The review demonstrates that combining content and collaborative information has increasingly replaced purely content-based or purely collaborative strategies.

B. Behavioural Signals and Their Semantics

A central challenge is that the same interaction may have different meanings. A repeated play can indicate strong preference, nostalgia, background listening, or simply an automatically repeated playlist. A skip can indicate dislike, contextual mismatch, or a desire to reach a specific portion of a playlist. Therefore, behavioural modelling should consider sequences and context rather than treating every interaction as an independent positive or negative label.

Listening frequency, recency, diversity of artists, genre concentration, new-release consumption, and session-level transitions can jointly characterize a listener. These characteristics can also be used to identify stable preference profiles and short-term changes. Spotify's generalized user-representation framework demonstrates the value of combining multiple signals and temporal scales within a shared embedding [3].

C. Temporal Behaviour

Music preference is dynamic. Long-term listening history may capture stable interests, whereas recent interactions can indicate temporary interests or changing preferences. A model that uses only long-term history may recommend familiar content while missing current intent. Conversely, a model based only on recent interactions may overreact to temporary behaviour.

A multi-temporal representation is therefore appropriate:

Long-term preference + medium-term transition + short-term intent → unified user representation.

Spotify's production-oriented framework explicitly aggregates interactions over approximately six months, one month, and one week, allowing the representation to retain stable interests while responding to recent activity [3].

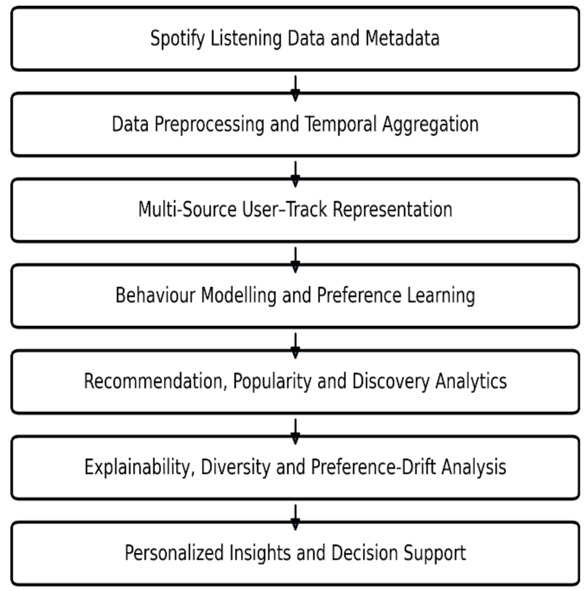

Fig. 1. Unified architecture for Spotify music-streaming behaviour analytics.

The proposed architecture in Fig. 1 places data acquisition before behavioural representation and separates representation from downstream intelligence. This structure prevents individual recommendation tasks from independently reconstructing the same user profile.

III. PERSONALIZATION AND MUSIC RECOMMENDATION

A. Evolution of Recommendation Approaches

Music recommender systems have historically employed content-based filtering, collaborative filtering, and hybrid approaches. Content-based methods recommend tracks with characteristics similar to items previously consumed by a user. Collaborative methods exploit similarities between users or item-consumption patterns. Hybrid methods combine both sources to overcome the limitations of either approach.

Content-driven recommendation has become increasingly important because music provides rich item-side information. Deldjoo et al. identify a transition from pure collaborative or content-based approaches toward models that combine collaborative and content information [2]. Such approaches are particularly useful for Spotify because track audio characteristics can complement behavioural information.

Deep learning further enables nonlinear user-item representations. Neural recommendation models, recurrent architectures, graph-based models, and transformer-based sequence models can represent relationships among users, tracks, artists, and listening sessions. However, increasing model complexity does not automatically resolve cold-start, diversity, explainability, or preference-drift problems.

B. Generalized User Representations

A significant development is the shift from task-specific features toward generalized user representations. Spotify's 2025 framework uses an autoencoder to compress rich multi-signal information into compact user embeddings, followed by lightweight downstream transfer models [3].

This architecture provides three advantages. First, the same representation can support several downstream tasks. Second, new features can be incorporated without manually redesigning every downstream model. Third, the representation can be updated more rapidly as user behaviour changes.

The reported Spotify framework also combines batch inference with near-real-time updates, allowing active users' representations to respond to new listening events while maintaining broader population coverage [3].

C. New-Release Recommendation

New music introduces a specialized recommendation problem. A user's general preferences do not necessarily predict how the user will respond to recently released music. Spotify's large-scale analysis found that new-release preferences differ from general music preferences and identified users who consistently prefer newer content [1].

This suggests that a comprehensive streaming analytics framework should include **content age** as an explicit behavioural variable. Instead of representing a user only by genre or artist preferences, the system should determine whether the listener tends to prefer recent releases, established tracks, older catalogue content, or a mixture.

IV. CONTEXTUAL, SOCIAL, AND DISCOVERY BEHAVIOUR

A. Context-Aware Listening

Context is fundamental to music consumption. A user may listen to energetic music during exercise, familiar tracks during routine activities, or calming music during relaxation. Consequently, identical users may exhibit different preferences under different conditions.

Context-aware recommendation seeks to incorporate variables such as time, session state, recent sequence, activity, and content age. The 2024 review of content-driven music recommendation identifies context awareness and sequence recommendation among the persistent challenges in the field [2].

Contextual models should therefore answer two questions:

1. What music does the user generally prefer?
2. What music is appropriate for the user's current situation?

The distinction is essential because a recommendation that is generally relevant may be inappropriate in a specific listening context.

B. Socially Motivated Music Discovery

Music also has a social dimension. Listeners may intentionally consume music that connects them with friends or communities. Spotify's socially motivated recommendation study examined a dataset containing listening activity from 100,000 users across 20 countries and modeled music that was trending within inferred communities [4].

The study introduced timeliness as an important consideration. A song may eventually be relevant to a user, but recommending it earlier can provide greater social value when the goal is to keep users informed about music circulating within their community [4].

These findings demonstrate that recommendation quality cannot always be represented by relevance alone. A comprehensive system should consider:

relevance + timeliness + social value + cultural context.

C. Diversity and Popularity Bias

A highly personalized recommender can repeatedly expose users to familiar artists and genres. This can improve short-term engagement while reducing discovery diversity. The broader recommender-system literature identifies popularity bias as a major concern because recommendation algorithms may disproportionately favor already popular items and reduce exposure for long-tail content [5].

In music, this issue is particularly important because the catalogue contains artists with vastly different levels of visibility. A system optimized only for immediate engagement may repeatedly recommend popular tracks rather than helping users discover emerging or less familiar artists.

Therefore, music recommendation should be evaluated using beyond-accuracy objectives such as:

- diversity,
- novelty,
- serendipity,
- catalogue coverage,
- artist exposure,
- and user satisfaction.

V. EXPLAINABILITY AND USER CONTROL

Explainability is increasingly important because music recommendation affects user discovery and listening habits. A recommendation that is relevant but incomprehensible may reduce users' confidence in the system.

Research on explainability in music recommender systems emphasizes that explanations should be evaluated separately from prediction accuracy and should account for the specific characteristics of music consumption [6].

Potential explanations include:

- "Recommended because you frequently listen to this artist."
- "Recommended because this track is similar to your recent listening."
- "Recommended because it matches your current discovery pattern."
- "Recommended because listeners with similar preferences enjoyed it."

However, explanation alone does not guarantee usefulness. Users may also require control over recommendation behaviour. Recent research involving music streaming and recommender-system interaction reports that users value controllability, diversity, and the ability to influence recommendation strategy, particularly when their listening goals are specific [7].

Thus, a future recommendation system should combine:
Recommendation → Explanation → User feedback → Adaptation.
This creates a closed interaction loop rather than a one-directional recommendation process.

VI. EVALUATION OF MUSIC-STREAMING ANALYTICS

Evaluation of music-streaming systems should extend beyond conventional accuracy. A recommendation can achieve high relevance while producing repetitive content, over-promoting popular artists, or failing for new users.

TABLE I
MUSIC-STREAMING TASKS, BEHAVIOURAL SIGNALS, AND EVALUATION DIMENSIONS

Task	Dominant behavioural signals	Suitable evaluation
User profiling	Artist, genre, recency, frequency	Profile consistency, clustering quality
Track recommendation	User–track interactions, audio similarity	Precision@K, Recall@K, NDCG@K
New-release discovery	Release age, recent listening, genre	Recall, precision, discovery rate
Sequential recommendation	Session sequence, recent tracks	Hit Rate, NDCG, next-item accuracy
Popularity analysis	Streams, artist, genre, temporal trends	MAE, RMSE, ranking correlation
Diversity-aware recommendation	Artist/genre/catalogue coverage	Diversity, novelty, coverage
Social recommendation	Community, country, genre, timing	Precision, timeliness, social value
Personalization	Multi-temporal user representation	Accuracy, cold-start gain, adaptation

A rigorous evaluation should also separate users and time periods appropriately. Randomly splitting interactions can allow information from the same user or temporal period to appear across training and testing sets, resulting in overly optimistic performance. Temporal evaluation is particularly important because real streaming systems must predict future behaviour rather than reconstruct historical interactions.
Spotify's generalized representation research illustrates the importance of both offline and online evaluation. Its reported framework was assessed through future-track prediction and large-scale online experiments, with improvements in discovery, consumption share, and search-related outcomes [3].

VII. UNIFIED SPOTIFY BEHAVIOUR ANALYTICS FRAMEWORK

Based on the reviewed literature, this paper proposes a conceptual **Unified Spotify Behaviour Analytics Framework (USBAF)**. The framework does not claim to represent an experimentally validated Spotify production system; rather, it synthesizes the major research directions identified in the literature into a unified architecture.
The framework contains five layers.

A. Data Acquisition Layer

This layer collects user listening histories, track metadata, artist information, audio characteristics, playlist interactions, temporal variables, release information, and where available, social or contextual signals.

B. Behaviour Representation Layer

The raw information is transformed into user, track, artist, and session representations. Multi-temporal aggregation separates stable preferences from recent transitions and short-term intent.

C. Behaviour Intelligence Layer

This layer performs:

- user clustering,
- preference prediction,
- recommendation,
- popularity analysis,
- new-release discovery,
- sequential prediction,
- and behavioural trend analysis.

Hybrid models combining content and collaborative information are particularly appropriate because music contains both rich item-side information and large-scale interaction signals [2].

D. Explainability and Adaptive Layer

This layer evaluates recommendation explanations, diversity, popularity bias, confidence, and user feedback. It also monitors preference changes and updates representations when the user's behaviour shifts.

E. Insight and Decision Layer

The final layer converts model outputs into actionable insights for listeners, artists, platform operators, and researchers. Listener-facing outputs may include personalized discovery, explanations, and preference summaries. Artist-facing insights may include emerging audience segments, release engagement, and regional listening patterns.

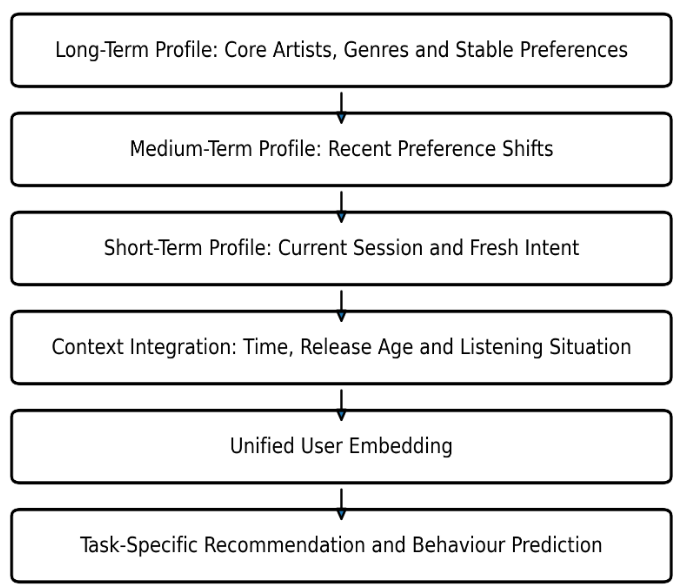

Fig. 2. Multi-temporal representation of music-streaming behaviour.

Fig. 2 emphasizes the importance of representing long-term, medium-term, and short-term behaviour simultaneously. This design is consistent with Spotify's recent generalized user representation work, which combines multiple temporal scales to capture stable interests and fresh intent [3].

TABLE II
RESEARCH GAPS AND DIRECTIONS FOR SPOTIFY-BASED BEHAVIOUR ANALYTICS

Existing limitation	Impact	Research direction
Single-task user modelling	Repeated feature engineering	Shared generalized user representations
Static preference profiles	Misses evolving interests	Multi-temporal and continual learning
Accuracy-focused recommendation	Repetitive recommendations	Diversity, novelty and coverage
Popularity bias	Reduced long-tail exposure	Exposure-aware recommendation
Weak context modelling	Contextually unsuitable tracks	Session and context-aware models
Limited explanation	Low transparency	Feature-, artist-, and behaviour-based explanations
Cold-start users	Poor early personalization	Onboarding and content-based representations
New releases treated as ordinary tracks	Weak emerging-music discovery	Dedicated new-release modelling
Limited social modelling	Missed community effects	Social and cultural recommendation
Offline-only evaluation	Uncertain real-world effectiveness	Temporal and online evaluation

VIII. OPEN CHALLENGES AND RESEARCH ROADMAP

Despite significant progress, several challenges remain unresolved.

1. Behavioural Data Sparsity

Not every listener generates the same amount of information. New users may have limited history, while inactive users provide sparse behavioural signals. Cold-start-aware representations are therefore essential.

2. Preference Drift

A listener's preferences evolve. Seasonal effects, changing interests, new artists, and social influences can shift listening patterns. Static profiles may therefore become outdated.

3. Popularity Bias

Popular artists may receive disproportionate exposure. Correcting this bias while maintaining recommendation quality remains challenging [5].

4. Diversity and Serendipity

A system should balance familiarity and discovery. Excessive similarity can create narrow listening patterns, whereas excessive diversity can reduce relevance.

5. Explainability

Users should understand why a track is recommended and, ideally, influence the recommendation strategy. Explanation quality must be evaluated independently from recommendation accuracy [6].

6. Cultural Differences

Listening motivations and social effects vary across cultures. Spotify's socially motivated recommendation study found differences in effectiveness across countries, emphasizing the need for culturally aware evaluation [4].

7. Real-Time Adaptation

Streaming systems must respond rapidly to new interactions. Spotify's recent work demonstrates a hybrid batch and near-real-time strategy for updating user representations [3].

8. Generative Recommendation

Recent Spotify research has begun exploring generative recommendation, including approaches for generating entire recommendation slates from natural-language prompts [8]. This direction could allow users to express high-level intentions such as a desired mood, activity, setting, or discovery goal rather than selecting explicit filters.

Research Roadmap

The research trajectory can therefore be organized into six stages:

Stage 1: standardized datasets and reproducible evaluation;
Stage 2: hybrid content–collaborative representations;
Stage 3: context-aware and multi-temporal personalization;
Stage 4: explainable, diverse and bias-aware recommendation;
Stage 5: continual and near-real-time preference adaptation;
Stage 6: generative and agentic recommendation capable of interpreting natural-language user intent.

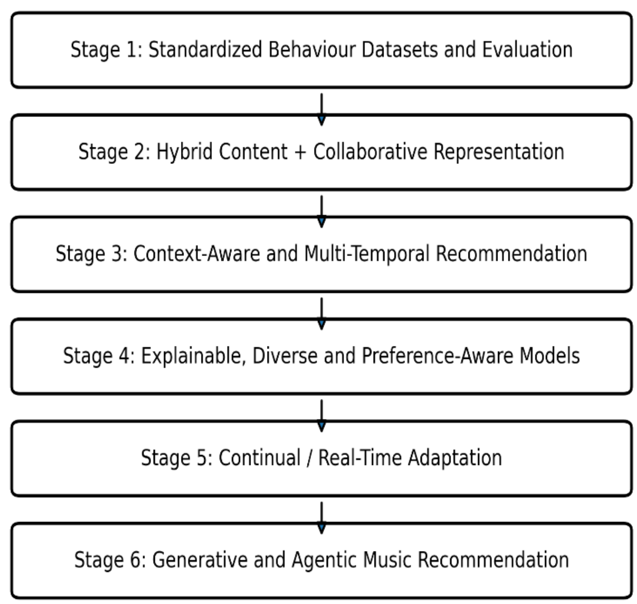

Fig. 3. Research roadmap toward adaptive music-streaming intelligence.

The roadmap moves from reliable behavioural data and standardized evaluation toward increasingly adaptive recommendation systems. The long-term objective is not simply a more accurate next-track predictor but an intelligent system capable of understanding changing user preferences, social context, discovery objectives, and recommendation constraints.

IX. DISCUSSION

The reviewed evidence indicates that music-streaming behaviour should be treated as a multidimensional and dynamic phenomenon. Audio characteristics remain important, but they are insufficient to explain listening behaviour independently. Collaborative interactions reveal community-level relationships, temporal signals capture changing preferences, and social variables explain forms of listening that are not purely individual.

A particularly important observation is that different recommendation objectives require different user representations. General music preference, new-release preference, social discovery, and short-term session continuation are related but distinct tasks. Spotify's research on newly released music demonstrates that users' preferences for new content differ from their general preferences [1]. Similarly, socially motivated recommendation requires attention to community trends and timeliness rather than individual relevance alone [4].

The emergence of generalized user representations provides a promising direction because a shared embedding can support multiple downstream applications. Spotify's 2025 work demonstrates the practical feasibility of multi-signal, multi-temporal user representations and reports improvements across discovery and search-related tasks [3].

Nevertheless, a universal music-behaviour representation should not optimize only engagement. A future framework should jointly consider relevance, diversity, novelty, social value, fairness, transparency, and user control. This is particularly important because recommender systems can influence the visibility of artists and the formation of listening habits.

The proposed USBAF therefore treats recommendation as only one component of a broader behavioural-intelligence system. The objective is to understand the user, represent the user, predict changing behaviour, explain recommendations, and continuously adapt the system while maintaining diversity and user control.

X. CONCLUSION

Music streaming has transformed music consumption into a continuously evolving interaction between listeners, content, algorithms, and social context. This review examined Spotify-related behavioural analytics across user representation, recommendation, new-release discovery, contextual listening, social recommendation, explainability, diversity, popularity bias, and adaptive personalization. The synthesis shows that isolated single-task models are insufficient for representing dynamic listening behaviour. A unified framework combining multi-source and multi-temporal representations, hybrid recommendation, contextual intelligence, explainability, diversity, and continual adaptation provides a stronger direction. The proposed USBAF architecture organizes these capabilities into five layers and establishes a research roadmap toward adaptive and generative music-streaming intelligence. Future research should prioritize standardized datasets, temporal evaluation, responsible personalization, culturally diverse evaluation, and user-controllable recommendation.

REFERENCES

[1] J. Jiang, A. Ponnada, A. Li, B. Lacker, and S. F. Way, "A genre-based analysis of new music streaming at scale," in *Proc. Web Science Conference*, 2024. Spotify Research reports the study using historical listening data from 1 million Spotify users and 282K newly released tracks.

[2] Y. Deldjoo, M. Schedl, and P. Knees, "Content-driven music recommendation: Evolution, state of the art, and challenges," *Computer Science Review*, vol. 51, art. no. 100618, 2024, doi: 10.1016/j.cosrev.2024.100618.

[3] G. Fazelnia, S. Gupta, C. Keum, M. Koh, T. Heath, G. Carrasco Hernández, S. Xie, N. Singh, I. Anderson, M. Hristakeva, P. P. Skidén, and M. Lalmas, "Generalized user representations for large-scale recommendations and downstream tasks," in *Proc. ACM RecSys*, 2025.

[4] B. Lacker and S. Way, "Socially-motivated music recommendation," in *Proc. ICWSM*, 2024.

[5] A. Klimashevskaia, D. Jannach, M. Elahi, and C. Trattner, "A survey on popularity bias in recommender systems," *User Modeling and User-Adapted Interaction*, vol. 34, pp. 1777–1834, 2024.

[6] M. Afchar, R. H. L. C. et al., "Explainability in music recommender systems," *AI Magazine*, 2022.

[7] "Music Recommendation Systems … and How They Want to Be Used," *International Journal of Human–Computer Interaction*, 2025.

[8] F. Tomasi, F. Fabbri, J. Carter, E. Kalomiris, M. Lalmas, and Z. Dai, "Prompt-to-Slate: Diffusion Models for Prompt-Conditioned Slate Generation," in *Proc. ACM RecSys*, 2025.

[9] Spotify Research, "Generalized user representations for large-scale recommendations," Spotify Research, 2025.

[10] Spotify Research, "How do people stream newly released music?" Spotify Research, 2024.

[11] Spotify Research, "Socially-motivated music recommendation," Spotify Research, 2024.

[12] Spotify Research, "Beyond the next track: Spotify Research at RecSys 2025," Spotify Research, 2025.