Feedback-Driven Gradual Discovery for Expanding Musical Preferences

Alec Nonnemaker
Delft University of Technology
Delft, Netherlands
alec.michael.n@gmail.com

Ralvi Isufaj
XITE
Amsterdam, Netherlands
ralvi@xite.com

Zoltán Szlávik
XITE
Amsterdam, Netherlands
zoltan@xite.com

Cynthia Liem
Delft University of Technology
Delft, Netherlands
C.C.S.Liem@tudelft.nl

Abstract

Many current recommender system techniques reinforce established tastes, leaving little room for venturing into unfamiliar music. A key challenge is our uncertainty about user preferences for previously unconsumed content, making it safer to build upon known preferences. To address this, we propose an incremental, feedback-driven method that gradually introduces users to new genres. By dynamically balancing recommendations between verified preferences and content with uncertain appeal, our approach maintains engagement while progressively expanding musical horizons. Adopting a Bayesian active learning approach, we update belief states iteratively as users provide feedback on new items. In a user study with data from a commercial music video platform, participants gradually discovered a previously unfamiliar music genre of their choosing. Comparing our method to both immediate genre introduction and passive small-step strategies without real-time adaptation, we observed significant improvements. Participants showed higher engagement with new music, stronger affinity for unfamiliar genres, and a greater sense of control, demonstrating the effectiveness of our iterative, feedback-informed strategy for broadening musical tastes. Supplementary code is available here[1].

CCS Concepts

• **Information systems → Recommender systems**; **Personalization**.

Keywords

Music Recommendation, Preference Expansion, Bayesian Active Learning

[1]https://github.com/alecmn/expanding_music_preferences

RecSys '25, Prague, Czech Republic

ACM ISBN 979-8-4007-1364-4/25/09
https://doi.org/10.1145/3705328.3748025

ACM Reference Format:
Alec Nonnemaker, Ralvi Isufaj, Zoltán Szlávik, and Cynthia Liem. 2025. Feedback-Driven Gradual Discovery for Expanding Musical Preferences. In *Proceedings of the Nineteenth ACM Conference on Recommender Systems (RecSys '25), September 22–26, 2025, Prague, Czech Republic.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3705328.3748025

1 Introduction

Modern recommender systems excel at delivering personalized content, yet discussions about their role in shaping user preferences remain contentious. While the concept of preference-reinforcing loops has been widely debated [1, 15], recent empirical studies indicate mixed and inconclusive evidence regarding their actual impact [3, 13, 17]. Nevertheless, within music recommendation there is a recognized need for systems that support "self-actualization", where systems help users develop and explore preferences beyond current tastes [5–7]. Current exploration approaches typically suggest representative songs directly from target genres [10]. However, when target genres differ significantly from current preferences, this will likely be too abrupt and overwhelming.

Therefore, in this paper, we address the research question: *How can we effectively guide a user towards developing a new taste based on their current preference profile?* The question of taste expansion can be considered as a unique cold-start variation [19]: we know current preferences, but not those yet-to-be-discovered. The challenge requires navigating multiple potential paths between initial preferences and target genres without knowing up-front which will be most effective. To enable these journeys, meaningful gradual transitions demand representations that capture nuanced relationships between songs across dimensions of similarity.

We investigate two hypotheses: **H1**) taking gradual steps from current preferences toward the target genre is more effective than immediately providing representative songs, and **H2**) integrating user feedback through Bayesian active learning can further improve effectiveness, balancing exploration and exploitation to gather preference information while maintaining satisfaction [22, 24]. Our research leverages data from an interactive music video platform with an extensive catalog and knowledge graph containing factual, expert-curated, and user-interaction data, providing an ideal testing ground for multi-modal, feedback-driven exploration approaches.

With this, we make two primary contributions: first, an adaptive learning approach to targeted music exploration that continuously refines recommendation paths based on user feedback, effectively addressing the critical need for frameworks that dynamically manage the exploration-exploitation tradeoff. Second, user study results demonstrating how this feedback-driven method guides listeners through personalized journeys into new musical territories while maintaining engagement, filling the gap in gradual discovery approaches that expand preferences without overwhelming users.

2 Related Work

We review relevant literature across two areas critical to our approach: preference elicitation techniques, and targeted exploration strategies in music recommendation.

Preference Elicitation for Novel Content. When guiding users toward genres they have not experienced before, user models inherently lack sufficient detail to provide accurate personalized recommendations. Traditional approaches to preference elicitation address this by separating exploration and exploitation phases, often reducing user enjoyment during initial interactions [20]. Research suggests that dynamically learning preferences through active learning offers advantages that allow the system to strategically select items that maximize information gain while maintaining user satisfaction [18]. Balancing exploration and exploitation can be achieved through techniques like Thompson Sampling [24], which adaptively adjusts the exploration-exploitation tradeoff as user feedback accumulates [12, 22]. Given the many potential exploration paths, efficient preference learning becomes crucial. Recent developments achieve this by sharing information across users with similar preferences through collaborative methods [26] or by leveraging similarity in embedded representation spaces to generalize feedback across related content [2, 25].

Targeted Music Exploration. To support self-actualization with recommender systems, several studies have examined targeted exploration scenarios where users actively seek to develop new music tastes. Liang et al. [10] investigated the effects of varying personalization levels when introducing users to new genres, finding that balancing personalization with representativeness improves exploration effectiveness. Their follow-up longitudinal study [9] demonstrated that user-controlled personalization adjustments enhance perceived system helpfulness. Taramigkou et al. [21] proposed taking gradual steps toward target genres by identifying paths through a user similarity graph. While this approach helped users understand how recommendations related to their tastes, it was limited by collaborative filtering constraints and lack of adaptive user control.

3 Data Representation

Music similarity is inherently complex and subjective [23], requiring representations that capture subtle transitions between musical styles for effective gradual exploration. While such representations are necessary for our approach, they are not our primary focus and remain independent from the exploration algorithm for easy replacement.

We represent music in a Collaborative Knowledge Graph (CKG) that fuses three signals: professionally curated metadata from a commercial streaming service (genres, editorial playlists), implicit user behaviour (likes, plays), and audio similarity. We add song-to-song edges derived from MULE embeddings [11]: a link is created when a song appears in another's top-10 nearest neighbours, weighted by their cosine similarity.

We embed this graph into a vector space using Node2vec [4], which balances computational efficiency with the ability to capture nuanced relationships. To evaluate our embeddings, we perform playlist completion on expert-curated playlists, and user-generated Spotify playlists [16], sampling 50% of songs as seeds and retrieving similar songs based on their embeddings.

Adding weighted audio similarity edges consistently improved embedding performance across both datasets. We experimented with the influence of these edges and report the top-performing configurations in Table 1 (we use the best-performing for the rest of this paper). The higher NDCG than Precision values indicate relevant songs are consistently ranked near the top. While performance on user-generated playlists is lower due to their diverse nature, the multimodal approach effectively captures nuanced song relationships, creating a representation space where similar songs cluster together while maintaining genre separation, which is critical for gradual exploration. Future work could explore more sophisticated representation techniques, though our results demonstrate that even with the selected representations, the adaptive learning approach significantly improves the exploration experience.

Graph	**Curated Playlists**		**Spotify Playlists**	
	Precision	NDCG	Precision	NDCG
Base CKG	0.4978	0.7488	0.0167	0.0963
+ Audio (0.5 weight)	0.5774	0.8631	0.0196	0.1101
+ Audio (0.25 weight)	**0.6455**	**0.9096**	**0.0200**	**0.1154**

Table 1: Playlist completion results comparing CKG embeddings with different audio similarity weights.

4 Gradual Discovery for Preference Expansion

To enable effective transition from users' current preferences to target genres, our approach utilizes an adaptive learning framework with two key components: path generation through high-dimensional spaces and feedback-driven recommendations that continuously refine based on user interaction.

4.1 Navigating the Representation Space

High-dimensional spaces present challenges for exploration (the "curse of dimensionality"). Rather than applying traditional dimensionality reduction, which degraded representation quality when evaluating on the playlist completion task, we construct a K-Nearest Neighbor (KNN) graph where each song connects to its 10 closest neighbors. This approach maintains the integrity of our original representations while enabling structured navigation.

To define starting and target points, we recognize that user preferences and target genres typically contain diverse subgroups. For example, a user might enjoy both Country and Hip-Hop, while the Hip-Hop genre itself contains distinct subgenres such as Southern Hip-Hop and Pop Rap. We identify central nodes representing subgroups for both user preferences and target genres using the PageRank algorithm [14], generating the starting set $S_{start} = \{s_1, ..., s_m\}$ and target set $S_{target} = \{t_1, ..., t_n\}$. To ensure diversity, each selected node is at least two hops removed, in order to capture distinct cluster centers. We then generate a set of standardized

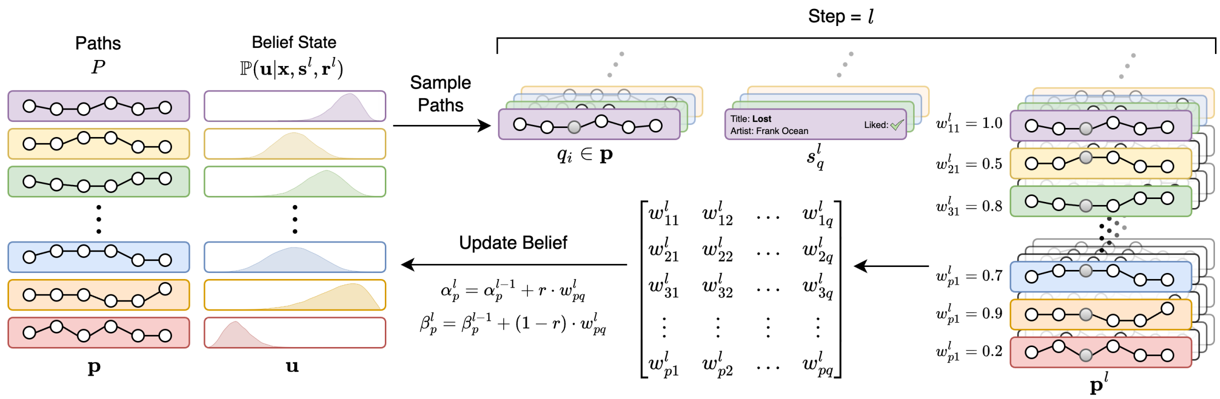

Figure 1: Our active learning approach: We maintain Bayesian beliefs over path utilities, sample paths for recommendation, and update all beliefs based on user feedback weighted by similarity to sampled paths.

paths $P = \{p(s,t) \mid s \in S_{start}, t \in S_{target}\}$, where each path $p(s,t) = [s, x_1, x_2, ..., x_l, t]$ has length L. These paths are created by finding shortest paths using Breadth-First Search and then standardizing their length through node insertion or removal.

4.2 Active Learning for Path Selection

Since we cannot predict which paths will maximize user satisfaction during exploration, we employ Bayesian optimization to dynamically balance exploration and exploitation of the paths (Figure 1). We maintain a belief state over each path's utility to the user.

Prior Beliefs: We initialize each path's utility u_p with a uniform Beta prior $\mathbb{P}(u_p) = \text{Beta}(1,1)$. The Beta distribution is used as it lies in $[0,1]$, aligning with binary user feedback. Additionally, since Beta is the conjugate prior of the Binomial distribution, it allows for straightforward posterior updates as new feedback is received.

Sampling and Recommendation: At each step, we sample from each path's posterior distribution following Thompson Sampling (TS) principles. Unlike traditional TS that would select just the highest-utility path, we select the top-20 paths with highest sampled utilities and present their corresponding songs at the current step.

Efficient Belief Updating: After receiving user feedback (like / dislike), traditional TS updates the belief for the sampled path: $\text{Beta}(\alpha_p^l + r, \beta_p^l + (1-r))$ where r is the reward (1 for like, 0 for dislike). To improve sample efficiency in limited-interaction scenarios, we update beliefs for all paths based on their similarity to the sampled paths. We introduce weight parameter w_{pq}^l, computed as the cosine similarity between songs s_p^l and s_q^l from paths p and q at step l. The updated posterior becomes $\text{Beta}(\alpha_p^l + r \cdot w_{pq}^l, \beta_p^l + (1-r) \cdot w_{pq}^l)$.

Our simulations showed that this approach significantly accelerates convergence to actual path utilities compared to traditional TS (Figure 2). This Bayesian active learning approach allows the proposed method to efficiently identify optimal paths through the representation space based on limited user feedback, balancing exploration of diverse paths with exploitation of paths already showing potential for user satisfaction.

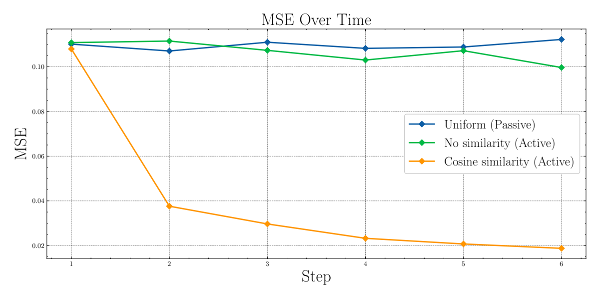

Figure 2: Comparison of efficient TS method using cosine similarity with traditional TS and a passive approach with no belief updating on a simulation exploration setting.

5 Evaluation

We conducted an online experiment to evaluate the effectiveness of our approach in guiding users toward developing new musical tastes. The experiment used a between-subjects design with three conditions to test our hypotheses:

- **Big Step (BS):** Recommendations immediately jump to the target genre and remain there for the duration of the experiment, serving as a baseline to evaluate the impact of gradual transitions.
- **Small Steps Passive (SSP):** Recommendations start close to user preferences and gradually move toward the target genre by randomly following generated paths, without incorporating feedback.
- **Small Steps Active (SSA):** Similar to SSP, but utilizes Bayesian active learning to incorporate user feedback and adaptively select paths.

Participants included employees from a music video platform (e.g., curation team members) and convenience-sampled individuals. After establishing preferences by selecting 20 enjoyed songs and a target genre, they completed 6 exploration steps of 20 randomly ordered recommendations at their own pace, within a period of 14 days with no per-day cap.

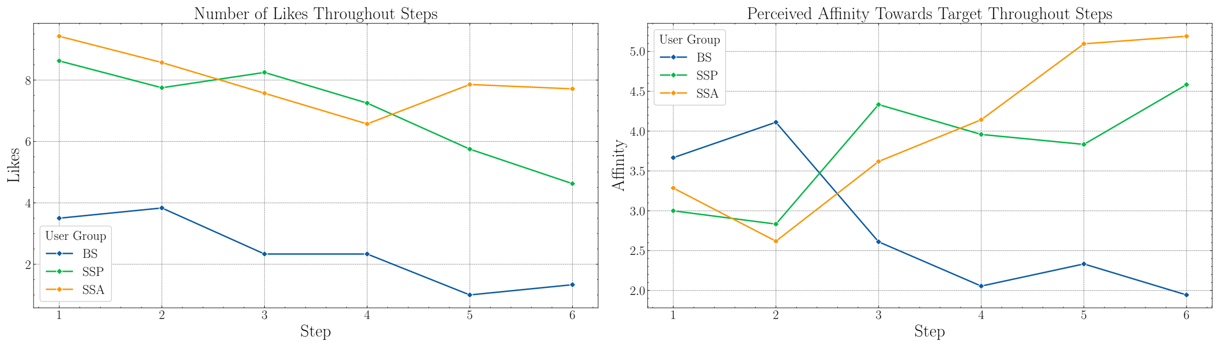

Figure 3: Progress of average *likes* and perceived affinity towards target genre throughout the steps for different user groups.

Regarding recruitment and sample size, 49 participants initially signed up, 30 began the study, and 21 completed it, resulting in a completion rate of 70%. Given the intensive nature of the study, requiring sustained engagement (average 1 hour per session) with multiple exploration steps and surveys, this completion rate and the final sample size represent substantial user commitment. Evaluating user experience over 120 songs yielded particularly rich interaction data from completing users, enabling comprehensive analysis of preference evolution and exploration patterns throughout the experiment.

The average age of participants was 28.1 years (SD = 6.43), with a gender distribution of 11 females and 10 males. Participants were randomly assigned to three groups: BS (N = 6), SSP (N = 8), and SSA (N = 7).

We measured effectiveness through both interaction metrics and user survey responses collected over time. For interaction data, we tracked the total number of songs liked and the number of target genre songs liked (indicating target affinity) throughout the exploration steps. User experience was assessed through 7-point Likert scales measuring perceived helpfulness of the approach at the end of the whole process and affinity toward the target genre after every step. Additionally, we collected subjective system aspects including perceived control, system understandability, quality of direction, and perceived personalization [8] after each of the six exploration steps. This evaluation framework allowed us to assess both hypotheses: (1) whether gradual steps are more effective than immediate target exposure, and (2) whether incorporating feedback improves effectiveness. By combining objective interaction data with subjective experience measures, we could evaluate the performance of our approach in facilitating personalized music discovery.

User Group	Total Likes		Target Likes/Song		Helpfulness		Final Affinity	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
BS	14.33	10.65	0.12	0.09	3.39	1.14	1.94	0.57
SSP + SSA	44.80	9.35	0.34	0.16	5.51	1.32	4.87	1.30
SSP	42.25	6.80	0.28	0.22	5.00	1.55	4.58	1.32
SSA	47.71	11.47	0.40	0.19	6.10	0.71	5.19	1.29

Table 2: Aggregated objective and subjective measures of effectiveness. We combine results from the SSP and SSA group to measure the isolated effect of 'gradual steps'.

6 Results and Discussion

Our online experiment gathered 21 valid complete responses across three conditions: Big Step (BS, N=6), Small Steps Passive (SSP, N=8), and Small Steps Active (SSA, N=7). User interactions and survey responses yielded several key insights.

6.1 User Engagement and Interaction Metrics

Our method significantly increased user engagement compared to the baseline approach (Table 2). Users in the gradual steps conditions (SSP+SSA) liked substantially more songs overall than the BS group (44.80 vs. 14.33, $\beta = 1.14$, $p < .001$). Moreover, they had a significantly higher ratio of target songs liked to target songs received (0.34 vs. 0.12, $\beta = 0.22$, $p = .02$), demonstrating a higher affinity towards the target genre.

Between SSP and SSA, incorporating feedback yielded modest improvements in both total likes (47.71 vs. 42.25) and target genre likes (0.40 vs. 0.28), though these differences were not statistically significant due to limited sample size. A temporal analysis reveals that although both groups' engagement initially decreased as recommendations moved away from initial preferences, the SSA group showed recovery in later steps, suggesting that feedback helped users discover enjoyable subgenres within their target (Figure 3).

6.2 Subjective User Experience

Users rated gradual approaches significantly higher than the baseline on all subjective measures. The SSP+SSA group reported higher perceived helpfulness (5.51 vs. 3.39, $\beta = 2.12$, $p = .003$), quality of direction (4.70 vs. 2.58, $p < .001$), personalization (4.40 vs. 2.38, $\beta = 2.02$, $p < .001$), control (4.36 vs. 2.15, $\beta = 0.99$, $p < .001$), and understandability (4.95 vs. 2.59, $\beta = 2.35$, $p < .001$).

Between SSP and SSA, incorporating feedback yielded notable improvements across all subjective measures, with the strongest effects on perceived understandability ($\beta = 1.04$, $p = .1$) and helpfulness ($\beta = 1.095$, $p = .11$). Mediation analysis revealed that improved personalization partially explained the impact of feedback on helpfulness ($\beta = 1.27$, $p = .082$).

Final affinity toward the target genre was substantially higher in the gradual approaches compared to the baseline (4.87 vs. 1.94, $\beta = 2.92$, $p < .001$), with personalization serving as a significant mediator ($\beta = 2.63$, $p < .001$). Temporal analysis showed that while BS users initially reported higher target affinity, this quickly declined, whereas SSP and SSA users showed steadily increasing affinity throughout the experiment (Figure 3).

7 Conclusions

Guiding users beyond the comfortable borders of their musical tastes requires a delicate balance between exploration and exploitation of preferences. Our experiment demonstrated that this balance is best achieved through gradual steps and continuous adaptation to user feedback. The results provide preliminary support for both research hypotheses: gradual exploration significantly outperformed direct genre introduction (**H1**), increasing engagement and target genre affinity, while feedback-driven adaptation further enhanced these benefits in later stages (**H2**).

The integration of real-time feedback helped users discover enjoyable subgenres within their exploration targets, addressing fundamental challenges in recommender systems by supporting self-actualization rather than merely reinforcing existing tastes. Despite limitations in sample size and experimental complexity, our results demonstrate the potential of feedback-driven approaches for preference expansion. An additional limitation to note is the absence of a control group receiving recommendations without explicit intent to guide users toward new genres, which may affect the interpretation of engagement-related outcomes. Future work should address these experimental limitations through simplified processes, inclusion of appropriate control groups, and platform integration, while extending this approach to other content domains and exploring more sophisticated representation techniques to further enhance exploration performance.

Finally, as for the impact of our work, while our approach demonstrates effective pathways for preference expansion, it does raise ethical considerations about agency and control. If target genres are chosen by platform operators rather than users, the system could potentially be repurposed to serve one-sided commercial interests, such as promoting sponsored content or driving traffic to higher-margin offerings rather than genuinely enriching user experiences. This tension between user autonomy and platform objectives underscores the importance of transparent implementation and clear user control in preference-expanding recommendation systems.

References

[1] Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor (Eds.). 2011. *Recommender Systems Handbook* (Boston, MA, 2011). Springer US. doi:10.1007/978-0-387-85820-3

[2] David Austin, Anton Korikov, Armin Toroghi, and Scott Sanner. 2024. Bayesian Optimization with LLM-Based Acquisition Functions for Natural Language Preference Elicitation. In *Proceedings of the 18th ACM Conference on Recommender Systems* (New York, NY, USA, 2024) *(RecSys '24)*. Association for Computing Machinery, 74–83. doi:10.1145/3640457.3688142

[3] Axel Bruns. 2019. *Are filter bubbles real?* John Wiley & Sons.

[4] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable Feature Learning for Networks. arXiv:1607.00653 [cs, stat] doi:10.48550/arXiv.1607.00653

[5] Nedim Karakayali, Burc Kostem, and Idil Galip. 2018. Recommendation Systems as Technologies of the Self: Algorithmic Control and the Formation of Music Taste. 35, 2 (2018), 3–24. doi:10.1177/0263276417722391 Publisher: SAGE Publications Ltd.

[6] Bart P. Knijnenburg, Saadhika Sivakumar, and Daricia Wilkinson. 2016. Recommender Systems for Self-Actualization. In *Proceedings of the 10th ACM Conference on Recommender Systems* (New York, NY, USA, 2016-09-07) *(RecSys '16)*. Association for Computing Machinery, 11–14. doi:10.1145/2959100.2959189

[7] Bart P. Knijnenburg, Martijn C. Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the user experience of recommender systems. 22, 4 (2012), 441–504. doi:10.1007/s11257-011-9118-4

[8] Yu Liang. 2023. Supporting Personalized Music Exploration through a Genre Exploration Recommender. ISBN: 9789038659015.

[9] Yu Liang and Martijn C. Willemsen. 2022. Exploring the longitudinal effects of nudging on users' music genre exploration behavior and listening preferences. In *Proceedings of the 16th ACM Conference on Recommender Systems* (New York, NY, USA, 2022-09-13) *(RecSys '22)*. Association for Computing Machinery, 3–13. doi:10.1145/3523227.3546772

[10] Yu Liang and Martijn C. Willemsen. 2019. Personalized Recommendations for Music Genre Exploration. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization* (New York, NY, USA, 2019-06-07) *(UMAP '19)*. Association for Computing Machinery, 276–284. doi:10.1145/3320435.3320455

[11] Matthew C. McCallum, Filip Korzeniowski, Sergio Oramas, Fabien Gouyon, and Andreas F. Ehmann. 2022. Supervised and Unsupervised Learning of Audio Representations for Music Understanding. arXiv:2210.03799 [cs, eess] http://arxiv.org/abs/2210.03799

[12] James McInerney, Benjamin Lacker, Samantha Hansen, Karl Higley, Hugues Bouchard, Alois Gruson, and Rishabh Mehrotra. 2018. Explore, exploit, and explain: personalizing explainable recommendations with bandits. In *Proceedings of the 12th ACM Conference on Recommender Systems* (New York, NY, USA, 2018-09-27) *(RecSys '18)*. Association for Computing Machinery, 31–39. doi:10.1145/3240323.3240354

[13] Lien Michiels, Jens Leysen, Annelien Smets, and Bart Goethals. 2022. What are filter bubbles really? A review of the conceptual and empirical work. In *Adjunct proceedings of the 30th ACM conference on user modeling, adaptation and personalization*. 274–279.

[14] Lawrence Page. 1999. *The PageRank citation ranking: Bringing order to the web.* Technical Report. Technical Report.

[15] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you.* penguin UK.

[16] Martin Pichl, Eva Zangerle, and Günther Specht. 2015. Towards a Context-Aware Music Recommendation Approach: What is Hidden in the Playlist Name?. In *2015 IEEE International Conference on Data Mining Workshop (ICDMW)* (2015-11). 1360–1365. doi:10.1109/ICDMW.2015.145 ISSN: 2375-9259.

[17] Amy Ross Arguedas, Craig Robertson, Richard Fletcher, and Rasmus Nielsen. 2022. Echo chambers, filter bubbles, and polarisation: A literature review. (2022).

[18] Neil Rubens, Mehdi Elahi, Masashi Sugiyama, and Dain Kaplan. 2015. Active Learning in Recommender Systems. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer US, 809–846. doi:10.1007/978-1-4899-7637-6_24

[19] Andrew I. Schein, Alexandrin Popescul, Lyle H. Ungar, and David M. Pennock. 2002. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* (New York, NY, USA, 2002-08-11) *(SIGIR '02)*. Association for Computing Machinery, 253–260. doi:10.1145/564376.564421

[20] Nicollas Silva, Thiago Silva, Heitor Werneck, Leonardo Rocha, and Adriano Pereira. 2023. User Cold-start Problem in Multi-armed Bandits: When the First Recommendations Guide the User's Experience. 1, 1 (2023), 2:1–2:24. doi:10.1145/3554819

[21] Maria Taramigkou, Efthimios Bothos, Konstantinos Christidis, Dimitris Apostolou, and Gregoris Mentzas. 2013. Escape the bubble: guided exploration of music preferences for serendipity and novelty. In *Proceedings of the 7th ACM conference on Recommender systems* (New York, NY, USA, 2013-10-12) *(RecSys '13)*. Association for Computing Machinery, 335–338. doi:10.1145/2507157.2507223

[22] Xinxi Wang, Yi Wang, David Hsu, and Ye Wang. 2014. Exploration in Interactive Personalized Music Recommendation: A Reinforcement Learning Approach. 11, 1 (2014), 7:1–7:22. doi:10.1145/2623372

[23] Geraint A. Wiggins. 2009. Semantic Gap?? Schemantic Schmap!! Methodological Considerations in the Scientific Study of Music. In *2009 11th IEEE International Symposium on Multimedia* (2009-12). 477–482. doi:10.1109/ISM.2009.36

[24] Xiaoxue Zhao, Weinan Zhang, and Jun Wang. 2013. Interactive collaborative filtering. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management* (New York, NY, USA, 2013) *(CIKM '13)*. Association for Computing Machinery, 1411–1420. doi:10.1145/2505515.2505690

[25] Sijin Zhou, Xinyi Dai, Haokun Chen, Weinan Zhang, Kan Ren, Ruiming Tang, Xiuqiang He, and Yong Yu. 2020. Interactive Recommender System via Knowledge Graph-enhanced Reinforcement Learning. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (New York, NY, USA, 2020-07-25) *(SIGIR '20)*. Association for Computing Machinery, 179–188. doi:10.1145/3397271.3401174

[26] Zhenyu Zhu, Liusheng Huang, and Hongli Xu. 2020. Collaborative Thompson Sampling. 25, 4 (2020), 1351–1363. doi:10.1007/s11036-019-01453-x