The Role of Serendipity in User-Curated Music Playlists

Valentina Maccatrozzo
v.maccatrozzo@vu.nl
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands

Tobias Kuhn
t.kuhn@vu.nl
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands

Davide Ceolin
davide.ceolin@cwi.nl
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands

Jacco van Ossenbruggen
jacco.van.ossenbruggen@vu.nl
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands

ABSTRACT

In this paper, we study the role of serendipity in music playlists. Serendipity is an important construct in recommendations, and finding an indicator of serendipity in a user-created playlist can facilitate the recommendation task. In particular, we want to know how the serendipity level of playlists is affected by the creator's ability and by the context they are created. To do so, we (1) measure the serendipity level of music playlists using a previously established Linked Open Data-based approach, (2) assess whether the ability of the creator of the playlists has an effect on the serendipity level, and (3) assess whether different contexts facilitate a higher or lower serendipity level of playlists. The serendipity level of playlists is calculated with the cosine distance between Linked Open Data Paths that connect the songs contained in the playlist. The ability of the creator to generate serendipitous recommendations is estimated by measuring his/her coping potential and assessing the genre diversity of listening history. We instrument a study using a Spotify playlists dataset. Previous results in different contexts suggest that the coping potential is a good proxy for the curiosity level of a person, and, in turn, for the diversified knowledge this person has. Our analyses confirm these findings also in the music context: we find that playlist creators with higher coping potential have a more diversified knowledge. They create a higher number of playlists that span across multiple contexts and genres. Conversely, a lower copying potential implies a lower number of less coherent playlists.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in interaction design**; *Empirical studies in HCI*; • **Information systems → Content ranking**; • **General and reference → Empirical studies**; **Metrics**.

KEYWORDS

Human-centered Data Science, User modelling, Music playlist Recommendations, Serendipity, Curiosity Theory

K-CAP '23, December 05–07, 2023, Pensacola, FL, USA

ACM ISBN 979-8-4007-0141-2/23/12.
https://doi.org/10.1145/3587259.3627552

ACM Reference Format:
Valentina Maccatrozzo, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. 2023. The Role of Serendipity in User-Curated Music Playlists. In *Knowledge Capture Conference 2023 (K-CAP '23), December 05–07, 2023, Pensacola, FL, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3587259.3627552

1 INTRODUCTION

Music streaming services have given users access to a possibly infinite collection of music, which, in turn, has generated the need for making this huge collection manageable. Playlists, automatically or manually generated, arose as the solution to this need. Personal playlists seem to be a way to satisfy the need for control of users [14].

Personal playlists are created for a variety of reasons. Some are made to contain most of the songs that a user likes (so-called playlist-container) or songs from the same artist, some are created for a specific reason, like a party or a workout, and some others are made to match a specific mood [49]. We refer to these as the playlist's context.

When creating contextualized playlists, users curate the content to their tastes and knowledge. While, according to them, this content is a good mix for the occasion, other users might find the combination serendipitous. This might be due to the unexpected context-song pair or to the memory-lane effect, i.e., when a song brings back memories [41]. This aspect is in contrast with the most traditional definition of serendipity in the recommender system community which quite often includes a component of novelty [15, 27, 30], and it aligns better with the more recent definition from de Rond [8]:

> *[...] serendipity results from the ability to identify 'matching pairs' of events, or events that are meaningfully, even if not necessarily causally, related. [...]*

This offers the opportunity to explore the concept of serendipity in the context of music playlist creation from a different perspective: while trying to create serendipity with a content-based method, we use a context-user-based method. In particular, we hypothesize that some users have the capability of creating the *matching pair* context-song(s), to which a listener might give a specific *meaning* and make it a serendipitous event. This new perspective on serendipity can be used to generate serendipitous music recommendations.

As a first step towards this direction, in this paper, we provide a deep analysis of contextualized playlists to lay the basis for a future recommendation algorithm. In particular, we start from the analysis of the serendipity level in different contexts: a playlist created for a party has the potential to be more serendipitous than a playlist created to focus. Hence, our first research question is:

RQ1: What is the level and distribution of serendipity in the different contexts of music listening, manifested as playlists?

In previous work, we proved that this capability to create/experience serendipity is well described by the curiosity level of a person, expressed in terms of coping potential [27, 28]. The coping potential is the ability of a person to cope with new knowledge. It reflects the person's openness to new, different experiences, and it has been widely used to understand, for instance, the ability to cope with visual art [42]. We hypothesise that this user's ability to cope with new knowledge could also indicate the ability to create serendipity. So our second research question is:

RQ2: What is the role of the creator's coping potential in the serendipity level of playlists?

When creating for a specific context, creators might feel inspired differently, and, hence, generate playlists with different serendipity levels. Hence, our last research question is:

RQ3: Do creators generate playlists with a similar serendipity level independently of the contexts?

To answer these research questions, we instrument an offline experiment on Spotify's playlists dataset [53]. The rest of the paper is organized as follows. We first introduce related work. In Section 3, we explain our approach, while in Section 4, we describe our experimental setup. Section 5 reports on our results, which are discussed in Section 6. We conclude in Section 7.

2 RELATED WORK

There is a limited availability of studies of serendipity in user-generated content, mainly due to the narrow interpretation of the word as a fortuitous event: it is perceived as an important aspect, but at the same time, researchers see it as unpredictable and not subject to control. At the same time, already Roberts discusses the "prepared mind in serendipitous scientific discoveries" [39], and more recently, de Rond analyses three such discoveries and shows that these were not pure accidents and that human abilities to understand such discoveries were pivotal in the discoveries themselves [8].

One of the first studies on serendipity in the music context [23] focuses on understanding why some users would experience the shuffle mode of the iPod as a serendipitous event through a qualitative analysis based on M&W's Threads of Experience [29]. They concluded that users who experienced serendipity were actively involved in understanding and interpreting the event.

Sun *et al.* [44] provide one of the first statistical analyses to measure unexpectedness, relevance and serendipity in micro-blogs. Besides proving how micro-blogs are a favourable environment to foster serendipity, they also show that once users experience it, they will be more engaged with the platform (i.e. produce more content). This is also one example of the ability of people to create serendipitous events for others, mainly involuntarily.

Hagen provides interesting insights about the generation of playlists, by analysing the behaviour of 12 heavy users of music streaming services [14]. Hagen performs an in-depth analysis of diaries about listening experiences, including date, location, and context. Additionally, users' social network activities were followed during the two months of reporting and face-to-face interviews were performed. The main findings are that users tend to create 3 types of playlists: static, dynamic and temporary. Static playlists once completed, never change, dynamic playlists include a constant increase in content, i.e. they become longer as time passes by, and temporary playlists are deleted as soon as the event for which they were created ends. Hagen also found that user-generated playlists often transcend standard classifications.

Krauser and North propose an analysis of contextualized music playlist listening [22]. By asking experiment participants to create playlists for eight everyday life contexts, they proved that the playlist's overall arousal reflects one of the contexts (e.g. playlists for jogging include high-arousal songs rather than a playlist for meditating). The analysis brings interesting results, however, it is not based on real-life user-generated playlists, especially because the contexts for the creation of the playlist were given.

Although most of the research on playlists focuses on recommendation algorithms to suggest the next track to add (see, for instance, [11, 13, 18, 46], and on the generation of complete playlists (see, for instance, [1, 4, 12, 16]), there are also studies similar to ours that perform an analysis to understand more the creation process. For instance, Ben-Elazar *et al.* focus on measuring the diversity of playlists measuring the tracks' acoustic differences, to understand users' different tastes for diversity [5]. Although they focus on diversity, this is partially similar to our analysis, but we focus more on song metadata.

Porcaro and Gómez [38] perform an analysis similar to ours, focusing on the diversity and popularity of the tracks in the playlists. They compare four playlist datasets: three user-generated and one with playlists from US radio stations. Their statistical analysis shows that user-generated playlists show higher ranges of diversity and popularity, especially for the more recent datasets. We continue on the same line of research, by adding content and context-related data and by assessing the serendipity of the playlists.

Jannach *et al.* analysed user-generated playlists and compared them with automatically generated ones [17]. They found that popularity, freshness and homogeneity are important features in user-generated playlists. They also found that these aspects are only partially reflected in the automatically generated playlists.

Zangerle *et al.* [53] perform a static analysis of user-generated playlists from the audio and lyrics features point of view. As audio features, they rely on the ones provided via Spotify's API[1]. While as lyrics features, they extracted acoustic, lexical, linguistic, semantic and syntactic features. They found that acoustic features represent the main characteristic that holds playlists together. [37] also analyse playlists concerning acoustic features confirming that users prefer different styles of music based on their mood or intended use. Continuing this work, Pichl and Zangerle [34] propose a multi-context-aware user model and track recommender system that jointly exploits information about the situation and musical preferences of users together with sound features to improve the recommendation task.

[1]https://developer.spotify.com/documentation/web-api/

Much research on serendipity and music is mainly in the context of music recommendations [24, 54] and music retrieval [40, 45]. Also, much focus has been on context for playlists and music listening in general [20, 33, 35, 47, 48, 51].

The problem of playlist context estimation has been addressed in different ways. For instance, Pichl *et al.* [35] estimate the context of playlists from their title, using lemmatisation and WordNet[2] to build a bag of words representing the context, and clustering these using K-means. However, they did not label these clusters, so we do not know what are these contexts for. While Chio *et al.* propose to semantically represent the context of a playlist with a low-dimensional embedding related to its title [7]. We propose a combination of these two methods to estimate the context of playlists, by using a context-labelled playlist titles dataset, created using Spotify's categories.

3 APPROACH

We aim to measure the serendipity level of contextualized playlists and analyze which are the variables that best contribute to this measure. First, we calculate the context of playlists and filter out those that fall outside specific contexts. Second, by adopting the SIRUP model presented in another work [27], we both calculate the serendipity level of the playlists and the coping potential of the playlists' creators (see Figure 1).

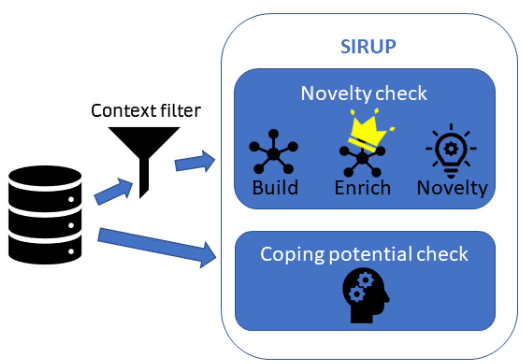


Figure 1: Serendipity Model

In this paper, we employ the SIRUP model in a different way than in previous work [27]. We use the novelty check to assess the potential serendipity level of a playlist, and, separately, we use the coping potential check to estimate the capability of the playlist creators. In this way, we aim to evaluate whether the coping potential is a good estimator of the serendipity capability across different contexts.

3.1 Context Filter

Estimating the context of playlists is a challenging task. Many playlists do not have context-related names, others are named with just one noun. Playlists' titles are very subjective so different users could use different nouns for similar contexts. For instance, when creating a playlist to be used during physical activity, one user could call it *Workout*, another could call it *Move*, and someone else could call it *Bike*.

To overcome this issue, we generate embeddings of the playlists' titles, to make them semantically more informative. Our approach is inspired by two other works, that also use the playlists' title to estimate the context [7, 36]. Word embeddings are vector semantics, used to represent a word as a point in a multidimensional semantic space that is derived from the distributions of word neighbour [19]. Word embeddings can be generated using trained language models (such as BERT [9]) so that words from the vocabulary are mapped to vectors of real numbers. Once the embeddings are generated, we can cluster them using a measure of similarity, such as cosine similarity. However, embedding-based clustering alone does not guarantee that the resulting groups of playlists correspond to different clusters. For this reason, in our experiment, we cluster playlists using a model trained on labelled playlist titles (for details, see Section 4).

3.2 SIRUP

The work presented in this paper is based on a model we presented in previous work of ours [27]. SIRUP is a recommender system model designed to generate serendipitous recommendations. This model is inspired by the curiosity theory developed by P. Silvia [43], and based on the fundamental theory of D. Berlyne [6]. Traditionally, curiosity is described as an internal conflict [6], that generates when there is a gap between the current knowledge level and the desired knowledge state. When this gap is perceived as manageable by the person curiosity gets triggered [25]. Silvia believes that curiosity depends on two processes: the novelty check of the item and the coping potential check [43]. Following this theory, we concluded that a recommendation is serendipitous when it generates curiosity in the person it is created for. In SIRUP, the novelty check assesses the item's novelty concerning the items in the user profile. While the coping potential check focuses on assessing the ability of the user to deal with such an amount of novelty. In this work, instead, the novelty check relates to the novelty of the combination of songs contained in a playlist and the coping potential check relates to the ability of a person to create such novel playlists: the higher the coping potential, the higher the ability to create a serendipitous playlist.

3.2.1 Novelty check. The novelty check aims at assessing the novelty of the combination of songs in a playlist. This is very similar to measuring the novelty of an item concerning a user profile in a content-based recommender algorithm. In this context, we define novelty as the opposite of similarity: if a song is less similar to tracks in the playlist, we consider it more novel.

As proposed in our previous work, [27, 28], we use Linked Open Data (LOD) paths to measure the novelty of a track in a playlist. Linked Open Data is the process of publishing structured data that allows for the enrichment and connection of the metadata, in such a way that different representations of the same entity can be connected [50]. A LOD path is an ordered set of types and properties, which connects two types, T_1 and T_{l+1}:

$$\{T_1, P_1, T_2, P_2, ..., T_l, P_l, T_{l+1}\}$$

where l is the length of the pattern [26]. To be able to extract these patterns, we first need a link between our items and a LOD dataset.

[2] https://wordnet.princeton.edu/

In our case, we perform semantic enrichment of the title of the track with DBpedia[3] concepts. Then, we extract patterns from the DBpedia knowledge space, between the aligned concepts.

Overall, the novelty check consists of the following steps:

(1) Playlist Graph Building: represent Spotify's information in a graph (RDF) format.
(2) Playlist Graph Enrichment: perform an alignment with DBpedia, to enrich the graph with extra information.
(3) LOD paths extraction: extract all the paths that exist between any pair of aligned DBpedia entities found.
(4) Serendipity calculation: average cosine similarity of all the patterns that link any relevant entities related to a given pair of songs.

Playlist Graph Building. For every playlist, we build an RDF graph with the information we collect through the Spotify API[4] for every song in the playlist: artist name, album name, and the release date of the album. As a schema, we use the DBpedia Ontology[5], given that the enrichment of the graph is performed with DBpedia (see next step).

Playlist Graph Enrichment. The enrichment of the playlist graph is performed with DBpedia. In particular, the enrichment process is performed as follows:

(1) Extract from the playlist graph the list of tracks, together with related artists and albums.
(2) Query the DBpedia SPARQL endpoint[6] incrementally:
  (a) query to find a musical artist in DBpedia with the name we have;
  (b) if a DBpedia URI for the artist is found, query to find an album of this specific artist with the title we have. If more than one artist is found, we identify the right one with this query, i.e., we keep the artist who has an album with the title we have. If an artist is identified, then we extract also the genres s/he plays;
  (c) if a DBpedia URI for the album is found, query to find a song in the album with the title we have. Additionally, we retrieve the genres and the categories of the album.

LOD Paths extraction. For all DBpedia URIs found (i.e., artist, song, and album), we perform path extraction. This step consists of executing SPARQL queries against the DBpedia SPARQL endpoint to extract all the existing paths between the aligned DBpedia entities, without performing a specific selection. For example, we can extract the connection between two songs that share the same recording location. In this way, we can find connections between tracks, which might explain the playlist-creation process. In this study, we focus on patterns of length 3 (i.e., composed of 3 properties). This choice is guided by the fact that patterns of length 2 (for instance, the path [song-artistName-song], where the property is "artist") are already included in the graphs and represent the standard metadata used in recommendation algorithms. We proved in [27] that serendipitous recommendations are fostered by leveraging similarity calculated with LOD paths instead of metadata.

Serendipity calculation. To calculate the serendipity of a playlist, we use the inverse of LOD paths' cosine similarity. In particular, we use the structure of these paths, i.e., the types and properties that compose them. We use the inverse cosine similarity measure (see Equation 1) for this purpose. The cosine similarity measure has several advantages, including the fact that it allows us to compare vectors of different lengths, which could happen in our case if we use LOD paths of different lengths.

$$inv\_cosine\_distance = (-1)\frac{\sum_{i=1}^{n} P_i \times T_i}{\sqrt{\sum_{i=1}^{n} P_i^2} \times \sqrt{\sum_{i=1}^{n} T_i^2}} \quad (1)$$

The similarity between two songs using LOD path components is calculated as follows. When there exist LOD paths (i in Eq. 1) that connect the two songs, we use the types (T in Eq. 1) and properties (P in Eq. 1) that constitute these paths as input for the inverse cosine similarity, like they are keywords that describe the songs. If there are no LOD paths that connect the two songs, their similarity is zero.

3.2.2 Coping potential check. The coping potential check aims at measuring the ability of the user to deal with new content. In our case, we use the coping potential as an indicator of knowledge: people with high coping potential have a higher curiosity level, so they have a more diversified knowledge than people with a lower coping potential. More diversified knowledge supposedly indicates a higher propensity to generate serendipitous playlists. Following the findings of our previous work [27, 28], we estimate the coping potential of playlists' creators as the number of genres of music they use in their playlists, as we show in Equation 2:

$$CP_c = \sum_{i=1}^{n} G_i \in \{P_c\} \quad (2)$$

where c indicates the creator, G the genre, and P_c the profile of the creator c.

4 EXPERIMENTAL SETUP

We use the ALF-200k dataset[7] collected by [53]. This dataset contains 17,889 playlists, 1,016 users and 671,672 unique tracks, including acoustic and lyrics features.

Playlist filtering is necessary to guarantee the uniformity of the dataset concerning playlist size. To decide on a good interval size, we analyse the quantiles of the lengths of the playlists in the original dataset and include in the dataset only playlists with lengths within the second and the third quantiles. This interval guarantees that the playlists are long enough to be informative and that we exclude playlist containers, i.e. playlists formed by everything the user likes. The quantiles of the lengths of the playlists are [1, 11, 16, 41, 10,051]. Hence, we included in the dataset those playlists of lengths between 11 and 41. After this filtering, the dataset contains 3,857 playlists, 770 users, and 114,080 unique tracks.

To perform the analysis we draw a sample from the dataset. The sampling has been performed based on the users. The sample is composed of 253 users (Confidence Level: 95%, Confidence Interval:

[3] DBpedia is a structured version of Wikipedia. More information at https://www.dbpedia.org/.
[4] https://developer.spotify.com/documentation/web-api/
[5] https://wiki.dbpedia.org/services-resources/ontology
[6] http://dbpedia.org/sparql
[7] Available at https://github.com/dbis-uibk/ALF200k.

5), 3,084 playlists (with an average length of 18.8), and 57,985 tracks (51,978 unique).

After the sampling, we estimate the context of the playlist. As explained in the previous section, we need a dataset to train a model to support cluster labelling.

Spotify uses a list of categories to tag items, including playlists. Example of categories' name are *Chill*, *Workout*, or *Gaming*, *Metal*, *Hip Hop* and *Jazz*. Not all categories' names indicate a context, but many do, so we consider only the ones that do. The data we used is anonymized for privacy reasons, so we cannot simply retrieve the category of the playlists in our dataset. Hence, we propose the following approach:

(1) Retrieve from Spotify the list of categories.
(2) Retrieve 50 playlist titles per category.
(3) Generate semantic embeddings per category using the playlists' titles.
(4) Generate embeddings per every playlist's title.
(5) Calculate the cosine similarity for all the pairs of playlist embedding-category embedding.
(6) The most similar category is selected as the playlists category.

Considering that we do not know where the creators of the playlist are from, and considering that some categories are country-specific, we collected categories through Spotify's API for 7 different countries that compose the majority of its users in 2017/2018[8] (Italy, Netherlands, Austria, Germany, France, UK and USA). In total, we collected 52 categories. On average, every category contains 59 playlists, with a minimum of 5 (*Afro*) to a maximum of 152 (*Trending* and *NL*). As explained in the introduction, we want to focus on playlists that do not belong to one music genre or one artist only. Hence, the contexts that we will focus on are: *At Home*, *Cooking & Dining*, *Focus*, *Fresh Finds*, *Gaming*, *In the Car*, *Kids & Family*, *Mood*, *Party*, *Pride*, *Sleep*, *Summer*, *Tastemakers*, *Travel*, *Wellness*, *Workout*. After estimating the context of playlists, we filter the dataset further: 887 playlists, 190 creators, and 15,764 unique tracks. After building the playlist graphs, we align the graph with DBpedia and enrich the graph with extra information, as explained in Section 3. We aligned a total of 13,834 entities, of which 686 songs. Between the aligned DBpedia entities, we extracted 320,131 LOD paths.

5 RESULTS

This section describes the analysis of the results obtained from the experiment described in the previous section.

5.1 Serendipity in Contextualized Playlists

We start by analysing the distribution of serendipity values in contextualized playlists. As we can see from Figure 2, there seems to be no difference in the distributions across different contexts, as confirmed by Wilcoxon tests.

[8] https://www.businessofapps.com/data/spotify-statistics/

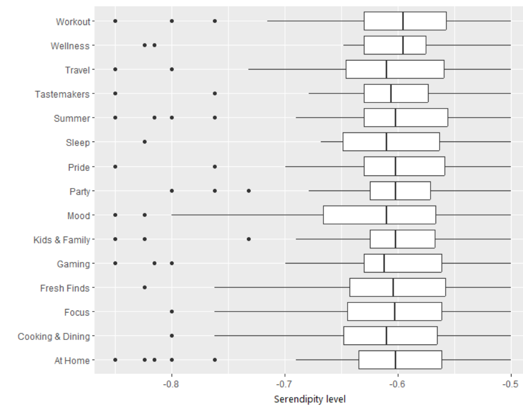


Figure 2: Boxplot of the distribution of the playlists' serendipity values grouped by context.

5.2 Playlist Serendipity and User Coping Potential

To analyse the relationship between the playlists' serendipity level and the creator's coping potential, we start with a correlation analysis. A Pearson correlation test is however not statistically significant (*p-value* = 0.13). We then perform a Wilcoxon signed rank test to compare the playlists' serendipity values grouping the data by coping potential ranges using the quantiles as a limit: this results in 4 coping potential groups ([0-22[, [22-40[, [40-71[, [71-77]). The only statistically different distribution is the first group (*p-value* = .002) proving that creators with lower coping potential create playlists with lower serendipity values. We additionally tested whether the 2 group distributions (using the second quantile as limit) come from the same distribution, using a Kolmogorov-Smirnov test (*D* = 0.11592, *p-value* = .03). With a *p-value* = < .05 we can reject the null hypothesis and say that the two sample datasets do not come from the same distribution.

A higher coping potential indicates a higher curiosity level, and, as a consequence, a more diversified knowledge. This is partially verified also in our data, as we find a weak correlation between the coping potential of the creator and the number of the created playlists (*cor* = 0.28, *p-value* = $6.737e^{-05}$) and with the average length of the created playlists (*cor* = 0.18, *p-value* = 0.01231). We also find a medium correlation (*cor* = 0.34, *p-value* < $2.2e^{-16}$) between the creator's coping potential and the number of contexts they create playlists for. We do not find a significant correlation between the coping potential of the creators and the serendipity level of the created playlists.

5.3 Context, Coping Potential and Playlists Serendipity

As we saw before, the context does not seem to influence the serendipity level of the playlists. We want to continue this analysis and see what changes if we take into account the creators' coping potential. First, we want to observe how the coping potential is distributed in different contexts. As we can see from Figure 3, there are some (statistically) significant differences. For instance, the context of *Tastemakers* tends to have creators with higher coping potential than the context *Pride* and *Party*.

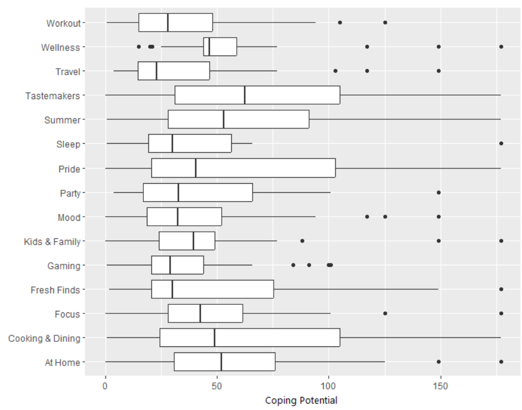


Figure 3: Boxplot of the distribution of the creators' coping potential grouped by context.

We also observe that, if we consider only creators with high coping potential (i.e. > 71) the distributions of the serendipity values in the contexts have more variance across contexts, and, specifically, tend to decrease. This is clear from Figure 4d, specifically looking at the context of *Gaming*. We perform a Wilcoxon signed rank test to confirm this observation, and we found that this is statistically proven only when compared with low coping potential users (W = 26810, *p-value* = 0.02431).

6 DISCUSSION

We perform an analysis of the ALF-200k dataset [53] to observe the role of serendipity in user-curated music playlists. Our first research question is:

> *RQ1: What is the level and distribution of serendipity in the different contexts of music listening, manifested as playlists?*

As seen in section 5.1 the level of serendipity across different contexts is very similar and does not distribute statistically differently. This could be expected as user-created content reflects many different tastes. Some contexts do have a higher variance. For instance the context 'Mood' has the higher variance in the sample: this is also well expected as Mood is a sort of context container for all sorts of moods, spanning from sad to exciting.

Then, we analyse the effect the creators' coping potential has on the creation process. In particular, we focus on the following research question:

> *RQ2: What is the role of the creator's coping potential in the serendipity level of playlists?*

To answer this question we study the distribution of the serendipity in the playlists created by the same user. Our result, indicate that creator with low coping potential seems to create playlists with statistically significant lower serendipity level. We also analysed the distribution of the playlist characteristics created by the same user and found the following:

- a weak correlation between the coping potential and the number of created playlists;
- a weak correlation between the coping potential and the length of the created playlists;
- a medium correlation between the coping potential and the number of contexts of the created playlists;
- no correlation between the coping potential and the serendipity value of the created playlists.

These results indicate that the coping potential is not a good proxy for serendipity generation in user-created playlists. The coping potential is the ability of a person to deal with new knowledge and can be used as a proxy for the curiosity level of this person [27, 42]. Our results confirm the fact that the higher the coping potential (i.e. higher curiosity) the more diverse the person's knowledge, in terms of length, number and contexts used. However, this diverse knowledge does not support the creation of serendipitous playlists.

Finally, we perform an analysis to study the effect the two factors (coping potential and context) have on the serendipity level of the playlists. These analyses aim to answer our third research question:

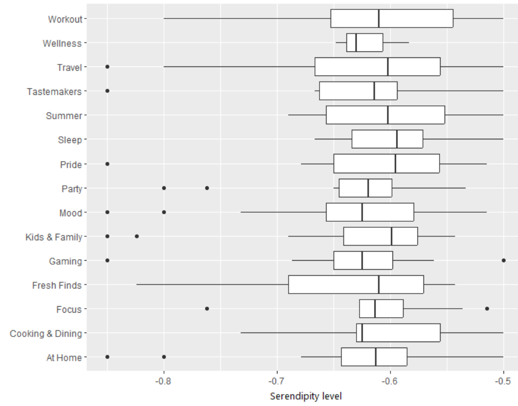
> *RQ3: Do creators generate playlists with a similar serendipity level independently of the contexts?*

There are some contexts where users with higher coping potential contribute more, like for instance in the context of *Tastemakers* and *Summer*. However, this is not reflected in the serendipity level of the created playlists, as clear from Figure 4. We can still make an interesting observation: in Figure 4d we can see that creators with higher coping potential seem to create playlists with lower serendipity levels. This is statistically proven when compared to users with low coping potential (see Figure 4a). It seems that high coping potential creators create a higher number of coherent playlists that span across multiple contexts and genres. Conversely, a lower copying potential implies a lower number of less coherent playlists.

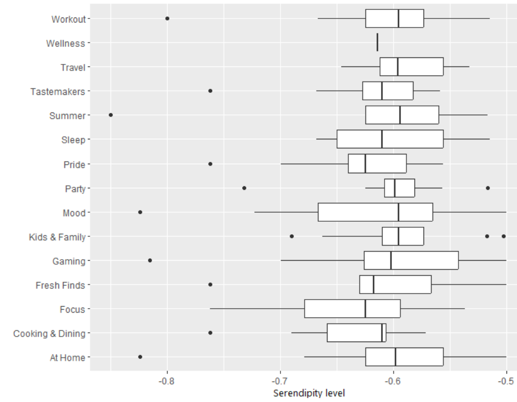
7 CONCLUSION

In this paper, we perform an analysis of user-created music playlists to study whether the coping potential of the creators and the context of creation have an effect on the creation process and, specifically, whether this allows the creation of serendipitous playlists. Being the coping potential a good proxy for the acceptance of serendipitous recommendations [27, 28], we hypothesize it could also be used as a proxy for the creation of serendipity. Our analyses suggest that (1) higher coping potential indicates higher diversified knowledge, (2) creators with high coping potential are more prolific in more contexts and (3) they tend to create more coherent playlists than creators with lower coping potential. These findings seem to indicate that creators with high coping potential use their diversified knowledge to organize their playlists in a coherent way in different contexts. While their counterparts, being less prolific and knowledgeable, organize their playlist in a less coherent way in less different contexts.

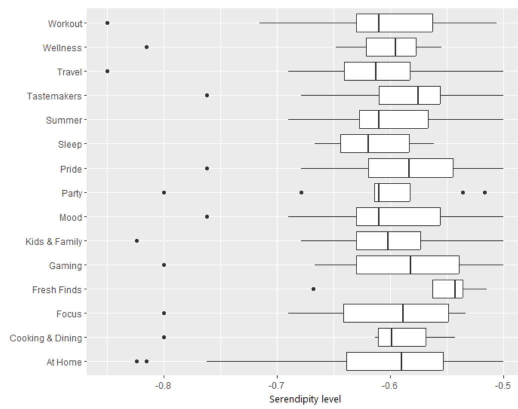
To the best of our knowledge, this is one of the first attempts to study the creation of music playlists with a user-centred perspective, as underlined by [10]. Works like [32, 37], focus more on song-based features, while, works like [14], focus more on the underlying motivation for making a playlist, rather than on the user factors that influence the creation, some others focus on the emotional



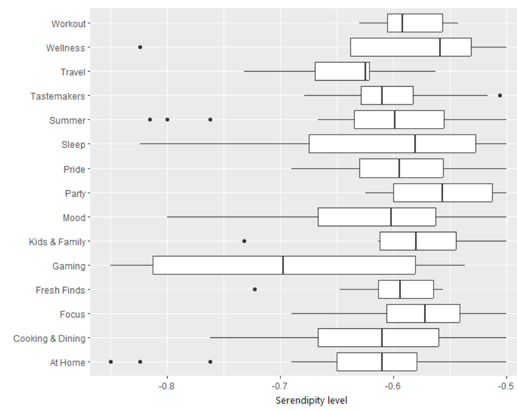
(a) low coping potential



(b) medium-low coping potential



(c) medium-high coping potential



(d) high coping potential

Figure 4: Boxplot of the distribution of the playlists' serendipity values grouped by different levels of coping potential.

aspects [3]. Some user-centred studies focus more on the music recommendation task, rather than on playlists [21]. Finally, some studies focus on the study of the influence of cultural background in music recommendations [52].

The findings of this work lay the basis for a human-centred approach to the understanding of playlist creation which accounts for the personal characteristics of the users that include not only tastes and preferences but also attitudes and personality traits. The results of such an analysis can be used to inform a recommendation algorithm to better support the user, for instance, in the playlist creation process or in the discovery of new genres. In this study, we focused on curiosity, but in the future, we would like to extend the study with a wider range of personality traits that might influence the creation process, in a similar fashion as done with music tastes [2, 31]. We are especially interested in how different personality traits affect the creation of playlists with different purposes/contexts.

REFERENCES

[1] Pedro Álvarez, Jorge García de Quirós, and Sandra Baldassarri. 2020. A Web System Based on Spotify for the automatic generation of affective playlists. In *Cloud Computing, Big Data & Emerging Topics*. Springer, Cham, 124–137.

[2] Ian Anderson, Santiago Gil, Clay Gibson, Scott Wolf, Will Shapiro, Oguz Semerci, and David M. Greenberg. 2021. "Just the Way You Are": Linking Music Listening on Spotify and Personality. *Social Psychological and Personality Science* 12, 4 (2021), 561–572.

[3] Willian G. Assuncao, Lara S.G. Piccolo, and Luciana A.M. Zaina. 2022. Considering emotions and contextual factors in music recommendation: a systematic literature review. *Multimedia Tools Applications* 81 (2022), 8367–8407.

[4] Mahta Bakhshizadeh, Ali Moeini, Mina Latifi, and Maryam T. Mahmoudi. 2019. Automated Mood Based Music Playlist Generation By Clustering The Audio Features. In *9th International Conference on Computer and Knowledge Engineering*. IEEE, 231–237.

[5] Shay Ben-Elazar, Gal Lavee, Noam Koenigstein, Oren Barkan, Hilik Berezin, Ulrich Paquet, and Tal Zaccai. 2017. Groove Radio: A Bayesian Hierarchical Model for Personalized Playlist Generation. In *Tenth ACM International Conference on Web Search and Data Mining* (Cambridge, United Kingdom). ACM, New York, USA, 445–453.

[6] Daniel E. Berlyne. 1954. A theory of human curiosity. *British Journal of Psychology. General Section* 45, 3 (1954), 180–191.

[7] Jeong Choi, Anis Khlif, and Elena Epure. 2020. Prediction of user listening contexts for music playlists. In *The 1st Workshop on NLP for Music and Audio*. ACL, 23–27.

[8] Mark de Rond. 2014. The structure of serendipity. *Culture and Organization* 20, 5 (2014), 342–358.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *2019 Conference of the North American Chapter of the ACL: Human Language Technologies*. ACL, Minneapolis, USA, 4171–4186.

[10] Ricardo Dias, Daniel Gonçalves, and Manuel J. Fonseca. 2017. From manual to assisted playlist creation: a survey. *Multimedia Tools Applications* 76 (2017), 14375–14403.

[11] Ricardo Dias, Joana Pinto, and Manuel J. Fonseca. 2014. Interactive Visualization for Music Rediscovery and Serendipity. In *28th International BCS Human Computer Interaction Conference on HCI* (Southport, UK). BCS, UK, 183–188.

[12] Marco Furini, Jessica Martini, and Manuela Montangero. 2019. Automated Generation of User-Tailored and Time-Sensitive Music Playlists. In *16th Annual Consumer Communications Networking Conference*. IEEE, Las Vegas, USA, 1–6.

[13] Anna Gatzioura, João Vinagre, Alípio M. Jorge, and Miquel Sànchez-Marrè. 2019. A Hybrid Recommender System for Improving Automatic Playlist Continuation. *Transactions on Knowledge and Data Engineering* 33 (2019), 1819–1830.

[14] Anja Nylund Hagen. 2015. The Playlist Experience: Personal Playlists in Music Streaming Services. *Popular Music and Society* 38, 5 (2015), 625–645.

[15] Leo Iaquinta, Marco De Gemmis, Pasquale Lops, Giovanni Semeraro, Michele Filannino, and Piero Molino. 2008. Introducing serendipity in a content-based recommender system. In *Proceedgins of the 8th International Conference on Hybrid Intelligent Systems*. IEEE, Barcelona, Spain, 168–173.

[16] Rosilde T. Irene, Clara Borrelli, Massimiliano Zanoni, Michele Buccoli, and Augusto Sarti. 2019. Automatic playlist generation using Convolutional Neural Networks and Recurrent Neural Networks. In *27th European Signal Processing Conference*. IEEE, Coruña, Spain, 1–5.

[17] Dietmar Jannach, Iman Kamehkhosh, and Geoffray Bonnin. 2014. Analyzing the characteristics of shared playlists for music recommendation. In *RSWeb Workshop at ACM RecSys '14*, Vol. 1271. CEUR, Silicon Valley, USA.

[18] Dietmar Jannach, Lukas Lerche, and Iman Kamehkhosh. 2015. Beyond Hitting the Hits: Generating Coherent Music Playlist Continuations with the Right Tracks. In *9th Conference on Recommender Systems* (Vienna, Austria). ACM, New York, USA, 187–194.

[19] Dan Jurafsky and James H. Martin. 2009. *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition*. Pearson Prentice Hall, Upper Saddle River, N.J.

[20] Peter Knees and Markus Schedl. 2013. A Survey of Music Similarity and Recommendation from Music Context Data. *ACM Trans. Multimedia Comput. Commun. Appl.* 10, 1, Article 2 (Dec. 2013), 21 pages.

[21] Peter Knees, Markus Schedl, Bruce Ferwerda, and Audrey Laplante. 2019. *9. User awareness in music recommender systems*. De Gruyter Oldenbourg, Berlin, 223–252.

[22] Amanda E. Krause and Adrian C. North. 2014. Contextualized music listening: playlists and the Mehrabian and Russell model. *Psychology of Well-Being* 4, 1 (2014), 22–38.

[23] Tuck W. Leong, Frank Vetere, and Steve Howard. 2005. The Serendipity Shuffle. In *17th Australia Conference on Computer-Human Interaction: Citizens Online* (Canberra, Australia). CHISIG Australia, Narrabundah, AUS, 1–4.

[24] Elad Liebman, Maytal Saar-Tsechansky, and Peter Stone. 2015. DJ-MC: A Reinforcement-Learning Agent for Music Playlist Recommendation. In *2015 International Conference on Autonomous Agents and Multiagent Systems* (Istanbul, Turkey). IFAAMAS, Richland, SC, 591–599.

[25] George Loewenstein. 1994. The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin* 116, 1 (1994), 75–98.

[26] Valentina Maccatrozzo, Lora Aroyo, and Willem Robert Van Hage. 2013. Crowd-sourced Evaluation of Semantic Patterns for Recommendation. In *UMAP Workshops*, Vol. 997. CEUR, Rome, Italy, 15–21.

[27] Valentina Maccatrozzo, Manon Terstall, Lora Aroyo, and Guus Schreiber. 2017. SIRUP: Serendipity In Recommendations via User Perceptions. In *22nd Intelligent User Interfaces Conference*. ACM, Limassol, Cyprus, 35–44.

[28] Valentina Maccatrozzo, Eveleine van Everdingen, Lora Aroyo, and Guus Schreiber. 2017. Everybody, More or Less, Likes Serendipity. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization* (Bratislava, Slovakia). ACM, New York, USA, 29–34.

[29] John McCarthy and Peter Wright. 2004. *Technology as Experience*. MIT Press, Cambridge, MA.

[30] Xi Niu, Fakhri Abbas, Mary Lou Maher, and Kazjon Grace. 2018. Surprise Me If You Can: Serendipity in Health Information. In *2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada). ACM, New York, USA, 1–12.

[31] Emily C. Nusbaum and Paul J. Silvia. 2011. Shivers and Timbres: Personality and the Experience of Chills From Music. *Social Psychological and Personality Science* 2, 2 (2011), 199–204.

[32] Savvas Petridis, Nediyana Daskalova, Sarah Mennicken, Samuel F Way, Paul Lamere, and Jennifer Thom. 2022. TastePaths: Enabling Deeper Exploration and Understanding of Personal Preferences in Recommender Systems. In *27th International Conference on Intelligent User Interfaces* (Helsinki, Finland). ACM, New York, USA, 120–133.

[33] Martin Pichl and Eva Zangerle. 2018. Latent Feature Combination for Multi-Context Music Recommendation. In *International Conference on Content-Based Multimedia Indexing*. IEEE, La Rochelle, France, 1–6.

[34] Martin Pichl and Eva Zangerle. 2021. User models for multi-context-aware music recommendation. *Multimedia Tools Applications* 80 (2021), 22509–22531.

[35] Martin Pichl, Eva Zangerle, and Gunther Specht. 2015. Towards a Context-Aware Music Recommendation Approach: What is Hidden in the Playlist Name?. In *International Conference on Data Mining Workshop*. IEEE, USA, 1360–1365.

[36] Martin Pichl, Eva Zangerle, and Gunther Specht. 2015. Towards a Context-Aware Music Recommendation Approach: What is Hidden in the Playlist Name?. In *International Conference on Data Mining Workshop*. IEEE, WashingtonUSA, 1360–1365.

[37] Martin Pichl, Eva Zangerle, and Gunther Specht. 2016. Understanding Playlist Creation on Music Streaming Platforms. In *International Symposium on Multimedia*. IEEE, San Jose, CA, 475–480.

[38] Lorenzo Porcaro and Emilia Gómez. 2019. 20 Years of Playlists: A Statistical Analysis on Popularity and Diversity. In *20th Conference of the International Society for Music Information Retrieval (ISMIR 2019)*. ISMIR, Delft, The Netherlands.

[39] Royston M. Roberts. 1989. *Serendipity: Accidental Discoveries in Science*. John Wiley & Sons, Inc., Hoboken, New Jersey, U.S.

[40] Markus Schedl, David Hauger, and Dominik Schnitzer. 2012. A Model for Serendipitous Music Retrieval. In *2nd Workshop on Context-Awareness in Retrieval and Recommendation* (Lisbon, Portugal). ACM, New York, USA, 10–13.

[41] Markus Schedl, Hamed Zamani, Ching-Wei Chen, Yashar Deldjoo, and Mehdi Elahi. 2018. Current challenges and visions in music recommender systems research. *International Journal of Multimedia Information Retrieval* 7 (2018), 95–116.

[42] Paul J. Silvia. 2005. Cognitive Appraisals and Interest in Visual Art: Exploring an Appraisal Theory of Aesthetic Emotions. *Empirical Studies of the Arts* 23, 2 (2005), 119–133.

[43] Paul J. Silvia. 2008. Interest - The curious emotion. *Current Directions in Psychological Science* 17, 1 (2008), 57–60.

[44] Tao Sun, Ming Zhang, and Qiaozhu Mei. 2013. Unexpected Relevance: An Empirical Study of Serendipity in Retweets. In *Seventh International Conference on Weblogs and Social Media*. The AAAI Press, WashingtonUSA, 592–601.

[45] 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 *7th Conference on Recommender Systems* (Hong Kong, China). ACM, New York, USA, 335–338.

[46] Andreu Vall, Matthias Dorfer, Hamid Eghbal-zadeh, Markus Schedl, Keki Burjorjee, and Gerhard Widmer. 2019. Feature-combination hybrid recommender systems for automated music playlist continuation. *User Modeling and User-Adapted Interaction* 29, 2 (01 Apr 2019), 527–572.

[47] Andreu Vall, Massimo Quadrana, Markus Schedl, Gerhard Widmer, and Paolo Cremonesi. 2017. The Importance of Song Context in Music Playlists. In *Poster Track of the 11th ACM Conference on Recommender Systems*, Vol. 1905. CEUR, Como, Italy.

[48] Mian Wang, Takahiro Kawamura, Yuichi Sei, Hiroyuki Nakagawa, Yasuyuki Tahara, and Akihiko Ohsuga. 2014. Context-Aware Music Recommendation with Serendipity Using Semantic Relations. In *Semantic Technology*. Springer, Cham, 17–32.

[49] Xinxi Wang, David Rosenblum, and Ye Wang. 2012. Context-aware Mobile Music Recommendation for Daily Activities. In *20th ACM International Conference on Multimedia* (Nara, Japan). ACM, New York, USA, 99–108.

[50] Liyang Yu. 2011. *Linked Open Data*. Springer, Berlin, Heidelberg, 409–466.

[51] Eva Zangerle and Martin Pichl. 2018. Content-based User Models: Modeling the Many Faces of Musical Preference. In *19th International Society for Music Information Retrieval Conference*. ISMIR, Paris, France, 709–716.

[52] Eva Zangerle, Martin Pichl, and Markus Schedl. 2020. User models for culture-aware music recommendation: fusing acoustic and cultural cues. *Transactions of the International Society for Music Information Retrieval* 3 (2020), 1–16. Issue 1.

[53] Eva Zangerle, Michael Tschuggnall, Stefan Wurzinger, and Günther Specht. 2018. ALF-200k: Towards Extensive Multimodal Analyses of Music Tracks and Playlists. In *Advances in Information Retrieval - 39th European Conference on IR Research, ECIR 2018*. Springer, Cham, 584–590.

[54] Yuan Cao Zhang, Diarmuid Ó Séaghdha, Daniele Quercia, and Tamas Jambor. 2012. Auralist: Introducing Serendipity into Music Recommendation. In *5th International Conference on Web Search and Data Mining* (Seattle, Washington, USA). ACM, New York, USA, 13–22.