Playlists and genre
McDonald, Callum; Foster, Allen; Rafferty, Pauline

Published in:
Journal of Documentation
DOI:
10.1108/JD-08-2023-0152
Publication date:
2024
Citation for published version (APA):

Document License
CC BY-NC

General rights
Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Download date: 02. Aug. 2024
Abstract

Purpose
Genre is a valuable access point for popular music collection, however, blurring of genre boundaries combined with changing listening habits and new forms of classification have brought genre’s importance into question. The playlist is now a common means of classification on Music Streaming Platforms. Recent commentary suggests that context is now a preferred access point. This exploratory study offers an examination of genre’s role in playlists.

Design/methodology/approach
A mixed methods study investigates, using Spotify, whether genre retains relevance amidst the rise in popularity of playlist-based music classification. Sample size is noted as a limitation of the study.

Findings
Qualitative coding of user and editorial playlist names revealed less than 20% of codes applied were genre-based. However, when non-genre themes were differentiated, genre themes ranked as one of the most prevalent. Context-based themes were most common, though genre was readily combined with other descriptive themes, highlighting its utility. Quantitative analysis of genre tags showed playlists with context-based themes demonstrated higher genre homogeneity than those using generic themes, indicating playlists were named on a form of genre-by-proxy basis.

Originality
The study suggests that genre continues to play an integral role in a field where an eclectic variety of descriptive themes has emerged, although its role may have changed. Context-based themes are central to the way users organise music, though such terms can often serve as containers for music collections sharing distinct generic and musicological similarities.

Keywords
Music Classification; Musical Genres; Spotify; Music Streaming

Introduction

Genre has traditionally been regarded as a valuable tool in the field of music information retrieval, helping to convey more of the ‘aboutness’ of a musical work, artist, record, or collection of works and typically surpassing musicological descriptions. For over half a century, within structures largely controlled by the record industry, genre has served as the primary means of engagement with popular music in many contexts including through radio stations, record stores, sales charts, critical commentary, and concert promotion. Challenges to the record industry’s control have emerged with the growth of Web 2.0-based services and the proliferation of digital accessibility across web-based platforms and portable technologies. In this context genre taxonomies have flourished, and arguably, weakened in authority as users and services describe growing music collections with greater freedom. ‘Streaming’ represents the latest content delivery method in the web-based
environment and allows large libraries of digital audio data to be delivered to the platform user for on demand and immediate playback (Hagen, 2015a, p. 13). Music streaming platforms (MSPs) also presents a new arena where both corporate and user interests can each exert a degree of curatorial control through the creation and management of playlists.

Recent studies have suggested that playlists are assuming a powerful role in the music industry, breaking the bounds of traditional generic classification (e.g. Prey, 2019, Prey et al, 2020, and see also, de Regt’s (2020) interesting commentary in which he refers to playlists as ‘the new genre’).

This paper offers an exploratory examination genre’s role in playlists by studying the naming and musical content of playlists on Spotify. Through a study of playlists belonging to ‘editorial’ (Spotify-curated) and user-curated groupings, it is hoped that light can be shed on the question of genre’s continuing role in modern music classification. In this study, popular music is understood to refer to music as that created with the intention of having a broad appeal and commercial success. It encompasses ‘rock, pop, soul, blues, country, jazz, gospel, as well as music from some world folk music traditions’ (Kishimoto and Snyder, 2016, p. 63).

‘Music genre’ and its application in knowledge organization

Genre has been described as exerting ‘enormous influence on music organization’ (Inskip, 2011, p. 3), conveying more of the ‘aboutness’ of music in few words than possible through musicological or bibliographic descriptions. While genre terms are powerful, harnessing their potential is a challenge as music genres form a complex network which is constantly changing form. Fabbri (1982) describes music genre as ‘a set of musical events [...] whose course is governed by a definite set of socially accepted rules’ (p. 52), which he divided into five distinct categories of formal and technical, semiotic, behaviour, social and ideological, and economical and juridical rules (pp. 54-59), highlighting the need for an interdisciplinary approach when considering music genre.

With the aid of Musicmap, Crauwell (2020) highlighted the problem of genres appearing, developing, and declining over time, while Zhang and Olson (2015) describe music genre as a collection of norms setting out the terms of engagement between creators and consumers of works. They use the notion of an essence-context duality to define genre – while genre possesses the ‘innate, immutable, independent of context’ properties of essence, it is made whole by context’s ‘fluidity of differing circumstances’, with these two characteristics traditionally presenting
difficulties for bibliographic control (p. 540). Music is well-suited to a genre classification, as ‘when used as a facet [...] genre enhances and extends the expressiveness of meaning’ (p. 551).

**Genre, the music industry, and technology**

Away from library classification, there has long been interest in the transformational power and influence of the music industry in relation to popular music (see, for example, Horkheimer and Adorno, 1944, Negus, 1997, Negus, 2011). The Frankfurt school’s somewhat pessimistic view, which distinguished between active agents of industry and the passive collective of consumers, established the notion of the culture industry, which, in and through technological affordances, standardizes and commodifies culture (e.g. Horkheimer and Adorno, 1944), turning it into a for-profit model. Technological and economic determinism underpin the logic of this view, but other scholars, perhaps more optimistically, note that the convergences between old and new media have made it possible for users to engage in user-generated content, participatory culture, and convergence culture, which offers possibilities of shifting and changing in relation to genre (see, for example, Connor and Katz, 2020).

Within this broader context, Frith (1996) sees genre as ‘a way of defining music in its market or, alternatively, the market in its music’, with generic decisions having a bearing on everything that happens to a performer or their work thereafter (p. 76). The creation of ‘genre worlds’ is brought into existence through a ‘loose agreement among musicians and fans, writers, and disc jockeys’ (p. 88). In a similar vein, Lena (2012, p. 118) portrays industry organisations as ‘classifying agents’, attempting to manage a disorganised field using simplified categories, while Negus (1999) relates notions of music genre as ‘a broader series of social divisions’ to ‘culture industry dynamics of musical production’ (p. 25).

Frith (1996) highlights the ‘seemingly inescapable use of generic categories in the organisation of popular culture’ (p. 75), arguing that while the music industry has always utilised labelling, its rules can be fuzzy. Genre serves as a means for the music industry to bring coherence to divisions across various media and market forces, but this approach struggles when consumer demand does not conform to generic classifications. In addition, genre, as a music industry tool, has changed over time as technology and commerce forced changes. Rossman (2012), for example, notes that radio stations began to differentiate in terms of content as technology became more accessible, giving rise to the ‘pop’ genre as a distinct broadcasting category (p. 66). Other genres emerged as a reaction to the ‘fluff’ of pop, as seen with the rise of ‘grunge’ (Lena, 2012, p. 22). However, even
genres with iconoclastic origins did not escape classificatory manipulation by the music industry. As grunge flourished, the music industry coined the term ‘alternative’ to refer to music of this type, creating an ecosystem spanning live concerts, television programming, and fashion (p. 24). Similarly, Frith (1996, pp. 84-85) details the creation, through agreement by committee of independent record label executives, of the ‘world music’ genre label in the late 1980s which encompassed ‘practically any music that isn’t [...] catered for by its own category’ with the aim of ‘improving the [...] sales situation’ of works of appeal to listeners of rock and folk music of disparate origin. Such commercial interests were at heart aimed at managing the relationship between publisher and listener.

The result of the music industry’s efforts to control a complex market was an expanding field of genre terms, exacerbated by lack of agreed categorical schema (Rossman, 2012, p. 66). For Pachet and Cazaly (2000, p. 2), the creation of taxonomies led to a convoluted landscape designed to create ‘the shortest possible path for consumers’ to music, and Aucouturier and Pachet (2003) articulated this through analysis of genre taxonomies on online music retail websites. They showed that within individual taxonomies, taxons did not bear fixed semantics, evidenced by Amazon’s use of taxons denoting period, country of origin, dance types, and more, and suggested that music genre was an ‘ill-defined notion [...] not formed on any intrinsic property of the music, but [depending] on cultural extrinsic habits’ (p. 84).

DiMaggio (1987) distinguished four main characteristics of ‘artistic classification systems’ based on differentiation, hierarchy, universality, and boundary strength (van Venrooij & Schmutz, 2015, pp. 799-800). Through these characteristics, and with regards to ‘the way in which artistic works are classified into ritually meaningful genres’ (DiMaggio, p. 446), it was posited that ‘the more differentiated the system of genre classification, the less universal’, in other words, agreement on classification schema reduces with increasing complexity, and this was echoed by Van Venrooij and Schmutz (2015), who noted that increasing genre diversity ‘signals changes in the institutional logic of the recording and radio industry that have focused on more and more specialized genre niches’, while the weakening of genre boundaries led to the rise in ‘hybrid’ genre forms as once-distinct commercial (typified by ‘frivolous’ pop music) and artistic (typified by ‘serious’ and ‘authentic’ rock music) values intertwine (p. 803).

Regev (2002), investigated ‘poprockization’ and the agglomeration of genre terms into nebulous families, arguing that while many genres exist, many of them are ‘interconnected in their histories
and stylistic genealogies’ (p. 252), particularly by a sharing of creative practices which Regev calls the ‘rock aesthetic’. Attempting to account for the rise in ‘pop/rock’ hybridisation, he employs Bourdieu’s ideas of ‘classificatory struggles’, claiming the institutionalisation of the ‘pop/rock’ entity ‘should be understood as corresponding to the emergence of new collective identities in the second half of the twentieth century’, (p. 260) specifically a highly technologically-literate ‘knowledge class’ consisting of ‘omnivorous’ high-status professionals, bound by various lifestyle identities, seeking legitimacy through engagement with contemporary cultural forms such as ‘pop/rock’ (pp. 261-262).

In relation to music and identity, the recent resurgence in vinyl as medium is perhaps of some interest, see, for example, Webster, 2020.

As taxonomies flourished, boundaries between genres blurred. DiMaggio suggests agreement on genre classification diminishes as differentiation increases, and Aucouturier and Pachet’s 2003 work indicates this holds true for music genre. New hybrid forms agglomerated based on aesthetic, practice, or socially driven factors developed. While not diminishing genre’s significance to popular music, these layers of complexity exacerbate difficulties in harnessing genre as a means of classification, at least when orchestrated by a music industry which plays an unending reactive role as it attempts to respond to metamorphic sociocultural forces. By the year 2000, popular music was associated with abundant genre terms describing ever expanding music collections. As differentiation threatened to undermine systems of genre classification, novel technologies and services threatened to weaken them further.

Community-driven folksonomies compare starkly with hierarchical taxonomies utilised by the music industry. Comparing the expert-made taxonomy of MP3.com and tags on Last.fm, Sordo et al. (2008) discovered the former type was better for hierarchical browsing, and the latter was suited to personal access, owing to a ‘flatter’ view of collections. They suggested a holistic approach to genre classification using ‘expert-based classifications, dynamic associations derived from the community driven annotations, and content-based analysis’ (p. 6). Santini (2011), comparing AllMusic.com and Last.fm, found strong superficial correlation between music genre taxonomies on each service, but this signal weakened as the level of detail used in analysis increased. Santini was wary of perceived pushback by tagging communities against the traditionally far-reaching influence of the music industry, causing these two groups to drift further apart at the expense of classificatory accuracy (p. 225).
As technology and economics have changed music distribution, the music industry first used genres, then created new genres, and then left genre with a reduced role, leaving us to ask if playlists are a means to document life experience, amplify moods, and describe collections of music in non-generic terms or merely a practical economically driven replacement ‘the new genre’ (de Regt, 2020). By examining the primary form of music distribution, the playlist, on what can be argued to be the most economically and culturally important medium of modern music communication, the MSP, greater understanding of genre’s role in popular music classification today can be attained.

**Music streaming platforms and playlist-based popular music classification**

Kusek and Leonhard (2005), frame music as a ‘utility’, likening an endless supply of digital music online, fed by increasing connectivity and portability of digital technology, to an ‘online music faucet’ (pp. 10-12). Drawing comparison to the rising popularity of radio, they envisage that ‘once wireless network access is an affordable and reliable standard […] digital music will take off and soar’ (pp. 14-15). The difference to the rise of radio is seen in an ‘increasing multitude of choices [outpacing] the single-minded purveyors of intellectual property’ (p. 15), resulting in services offering increasing personalisation, social connectivity, and recommendation-based listening opportunities (p. 27).

Morris (2015a, p. 175) considers what is to become of listening habits when music as a commodity becomes increasingly pervasive in everyday life across technologies and services, citing Kassabian’s ideas of ‘ubiquitous listening’ (Kassabian, 2001, p. 16 in Morris, 2015a, p. 175) where listeners engage in other activities, creating a ‘sourcelessness’ that at one time was associated with public spaces, but is now part of the private listening space. When Huber (2019) compared two surveys of 1,000 people to study public attitudes to music, results showed little difference in concert attendance or amount of incidental music listening, but the ‘Web 2.0 generation’ showed more awareness of how they used music. Over two thirds of this group rated music’s ability to create the right atmosphere, while only one third of older people shared this sentiment. Similar differences in the perceived importance of music as an atmospheric canvas for social events, as a reminder of experiences or people, or as a means to bond with friends, tying directly into Kusek and Leonhard’s notions of music as a utility.

Mazierska et al. (2019) see the launch of YouTube and Spotify as the beginning of a period of ‘advanced convergent digitisation’, where ‘streaming supplanted other types of music consumption as consumers needed to pay nothing or very little to access […] music,’ (p. 6). The result of this
period, and the associated drop in importance of physical records, entails what the writers call the
‘unbundling’ of the album and rise in use of playlists. With fragmentation of published music
collections through this ‘unbundling’, combined with the use of playlists working in tandem to
individualise the listening experience, Morris (2015a) concludes that ‘context has usurped content’
(p. 188). However few studies have focused on playlist-based music classification while those that
have focus on the music industry, artists and user behaviour relating to playlists, such as Johnson
(2018).

In 2013, Spotify shifted to personalised playlists to target ‘every mood and moment’ (Spotify, 2013
in Eriksson et al., 2019, p. 138). Playlists play a central role in attracting advertisers by displaying
relevant adverts (Eriksson et al, 2019, p.160). Pierce (2017) highlighted the impact of playlists on
music promotion, citing a data-driven approach that led to a song’s success through placement on
popular playlists. Aguiar and Waldfogel (2018) found that being added this playlist resulted in 20
million additional streams and increased revenue of up to US$163,000, while analysis of Spotify’s
Twitter account shows Spotify actively promoting of playlists since 2013 all promotion of albums
and individual tracks (Prey et al., 2022, p. 80).

As Eriksson (2020) notes, playlists as a means of organising digital media have been in use for
decades, but they have only recently become a central means of music promotion, with Spotify
leading the way among MSPs in this regard (p. 415). Eriksson focuses on playlists as a ‘container
technology’ (a means of enabling music transportation), facilitating operation of the online music
economy. She expresses concern over abuse of the playlist system, whether through bias towards
the selection of tracks by artists signed to major record labels on curatorial playlists, or the use of
‘bots’ to manipulate the quantity of playlist interaction (pp. 422-423).

Playlists can have significant sway on listening habits, enabling songs to accrue millions of listens,
provided they find their way onto the right playlists at the right time, in a process largely
coordinated through analytics data on Spotify. Having already moved away from album-based forms
of distribution, generic organisation also appears to be at risk. Quoting Spotify CEO Daniel Ek, ‘we
want to soundtrack every moment of your life […] it’s really all about bringing music to more
moments in your life’ (Pierce, 2017). These comments are worth bearing in mind as user-focused
studies of playlist consumption are examined.
**User-focused studies of playlists**

As one of the first works on user behaviour relating to playlists, Hagen’s (2015b) study focused on the pre-digital era by necessity, though it nonetheless has worthwhile findings. She paid particular attention to the ideas of Benjamin, Baudrillard and Shuker who each consider the ritual action of collection, though Hagen focussed on a format that has ‘apparently surrendered its physical materiality: music’ (p. 627). She considers the implications of physical formats on listening and storage preferences, and the disruption of digital formats on album based presentations by enabling the listener to select tracks at their leisure. Burkart (2008, p. 247 in Hagen, p. 627) asserts that digital formats make it impossible to ‘collect’ music as such, necessitating the use of ‘symbolic substitutes’.

Using a self-reported diary study, Hagen (2015b) recruited heavy users of MSPs and followed their Facebook profiles and ‘scrobbling’ activity, revealing several behaviours relating to the curation of playlists. These included highly individualised motives for playlist organisation, with one user having playlists sorted by the tuning of individual tracks to use when practicing bass guitar, and another having a playlist titled ‘Unique’ within which the vocals qualities of musicians stood out from the ordinary (pp. 635-636). There were also a diverse collection of context-sensitive playlists and associated behaviours. Users created temporary playlists to fit a given mood in the moment, or regularly updated a select playlist to reflect their changing tastes in music over time. Playlists were used in relation to the self and social situations, in some cases important life events, offering space for reflection and reminiscence (p. 637). Hagen suggests these behaviours indicate a desire to replicate rituals of physical collecting even in subscription based services where users pay for the privilege of access to their collections (pp. 643-644).

Siles et al. (2019) examined notions of playlists as ‘affective genres’ or, in other words, ‘fusions of musical substance, sociotechnical assemblages, and socio-material practices that respond to the exigencies of affect’ (p. 1). Interviews with MSP users showed they were appropriating playlist-based curation in order to ‘cultivate’ affect. This was done by creating moods in both individual and social contexts via collection of certain types of music, and also as a response to moods, with users turning to playlists to either sustain moods or reflect moods felt in a given moment (p. 4). Playlists serve as ‘affective genres’, first through the playlist’s origin in an affective exigence which is then reconciled ‘with specific musical substance’ (p. 5) to create what is, in effect, a ‘genre’. Siles et al. consider music genres as a relatively pre-defined set of musical properties, likening them to the efforts of users in establishing similarities between the songs in their playlists. They report on users
often making references to playlists with ‘rock’ or ‘jazz’ feelings, with these definitions ‘honouring a pre-existing symbolic contract’, usually meeting ‘the expectations about the kind of music they contain’ (p. 5).

Besseny (2020) notes that Spotify lacks major folksonomy-friendly functions such as tag clouds, instead utilising a series of menus and a series of featured content pages. ‘The immensity of music genre’ (p. 11) is indicative of the problems MSPs face in integrating effective wayfinding features, with the task of mapping every single genre on the service to every song with tagging being an impractically massive undertaking. While a search function exists, it is limited to the extent of the user requiring some knowledge of what they are searching for. These factors amalgamate to draw users towards playlists, either through accessing Spotify’s collection of editorial playlists or creating their own. Besseny describes playlists as ‘Spotify’s main area for user-created and user-curated content’, and playlist naming is one of the key folksonomic features of the platform (p. 13). This is a significant contrast to earlier folksonomy-driven services, such as Last.fm.

These studies show users as utilising playlists for very personal means, in contrast to MSPs’ own usage which is centred around the data-driven logistics of music delivery. In relation to MSPs, playlists now serve as practically the only interface for users to exert a degree of curatorial control, and they are crucial in that they allow users to replicate collecting rituals in a subscription-based space that denies them ownership over their music collections. Thus, assisted by the increasing pervasiveness of music in everyday life, users have assumed a diverse collection of behaviours relating to playlists, particularly surrounding their use in utility and mood-based contexts, seeing a significant departure from generic forms which were a mainstay of music classification for over half a century.

**Methodology**

The methods in this study support a qualitative assessment of editorial and user-made playlists, allowing sub-samples to be identified and examined within genre and playlists, and facilitating quantitative analysis of musicological elements. Spotify was chosen as a data source for several reasons: its presence as a touchstone in discussions of the rising influence of MSPs, its status as the largest MSP based on subscription market share (Mulligan, 2022), and accessibility to assorted playlist data. Moreover, the majority of MSP-based studies make explicit reference to Spotify (see Johnson (2018), Webster (2021), Hracs and Webster (2021), Hagan (2015b) and Siles et al. (2019)).
These works suggest editorial and user playlists are of sufficient importance to warrant investigation. While popular music genre is often charted based on cultural rather than musicological considerations, the work of Abrahamsen (2003) provided inspiration for investigation of musicological properties of songs within playlists.

**Research aim**

The aim of this research is to examine the role of genre in playlist-based music classification. The research task is two-fold. First, this works seeks to understand the role of music genre in naming of edited and user-generated Spotify playlists. To this end, an analysis of Spotify playlist names was conducted for thematic categories of playlist naming conventions. The research questions for this first research task are posed as follows.

- **RQ1.1** What are the most frequent naming convention types for edited and user-created Spotify playlists?
- **RQ1.2** What is the role of music genre terms in the naming of these Spotify playlists?

After the naming conventions of Spotify playlists have been observed, this work investigates the qualities of the songs in the edited playlists representing different naming conventions. Namely, the homogeneity of the songs’ tag and unigram-based classifications and their musicological properties.

For this second research task, the research questions are posed as follows:

- **RQ2.1** To what extent are the songs in edited Spotify playlists representing different naming conventions homogeneous in their tag and unigram-based classifications?
- **RQ2.2** To what extent are the songs in edited Spotify playlists representing different naming conventions homogeneous in their musicological properties?

**Sampling the playlists**

The Spotify Browse interface serves as a primary means of music discovery on Spotify. Selecting this option via the platform’s web player presents the user with an interface featuring a search bar and a series of tiles underneath a ‘Browse all’ heading. If the user opts not to use the search function, these tiles serve as the start of the browsing process for editorial playlists. These ‘parent categories’ use a diverse vocabulary based on genre, context, and geographical terms.

After navigating into a parent category, a series of sub-categories are presented, with playlists separated into various categories such as ‘Popular Playlists’ or ‘Summer Sounds’. In some cases, a deeper level of further sub-categories may be presented offering an even more granular selection of
playlist choices. Playlists serve as individual units of analysis for the sake of this study. Categories containing only podcasts, songs or albums were ignored. Where categories contained a mixture of playlists, songs, albums, and podcasts, only categories containing a minimum of two music-based playlists were included in data collection.

Data collection began on Monday, August 15, 2022. As the full data collection process would take longer than a single day, screen captures of all relevant web pages were created using the GoFullPage browser extension for Google Chrome (Full Page, 2022). This ensured a ‘snapshot’ of playlist data could be acquired from a fixed point in time (though even the screen capture process was approximately five hours) to mitigate against the risk of Spotify altering its editorial playlist offering during data capture.

Following this, category and playlist names were transcribed into a Microsoft Excel spreadsheet. In total, 4,180 playlists were tabulated, with each playlist being allocated a row in the data table, and given a unique ID value. This value corresponded to the order in which playlists are presented in the Browse interface – first by parent category, then by any subcategories, and finally by the order of playlists contained within any individual category.

As user playlists lack a structured classification system, devising a sampling method to acquire playlists from the approximately four billion available on the platform (Spotify, 2021) proved to be challenging. Initially, sampling was intended to use an MSP analytics service such as Chartmetric (2022a). However, such services prove unable to generate a ranking of playlists created by ‘regular’ users independently of commercial creators., and even a ranking of ‘Independent Curator’ playlists contained playlists created by industry figures (Chartmetric, 2022b).

To avoid this problem, it was decided to make use of Spotify’s Million Playlist Dataset. The dataset was made publicly available in 2020 (Chen, 2020), and contains a random sample of a million playlists created by Spotify users based in the USA. Playlists are anonymised, were public at the time of the creation of the dataset, contain between 5 and 250 tracks, at least three unique artists and two unique albums, and were created between January 1, 2010 and December 1, 2017 (Spotify, 2020). While this dataset lacked the recency of the editorial playlist data captured for this study, it was considered to be the most representative user playlist data available at the time of the study.

The dataset was supplied as JSON files stratified into 1,000 segments of 1,000 playlists each. Based on the length of time taken to code editorial playlists, an optimised sample size of 1,000 playlists,
equivalent to a single JSON file in the dataset, was selected. As playlists for the dataset were sampled randomly, it was decided that selection of a single JSON file from the thousand available would be a sufficiently randomised data sampling method. The JSON file was then imported into Microsoft Excel and converted to a spreadsheet format using the Power Query feature.

Content analysis of playlists entailed a process of coding which qualitatively assessed and categorised playlist names and quantified instances of each form of code. After the initial coding pass, eleven facets were established: Genre, Context, Musical Property, Non-Musical Sound, Geographic Origin, Popularity, Temporal Quality, Abstract Quality, Branding/Promotional, and Named Artists, and a catch-all ‘Other’ facet. A second pass involved closer inspection of playlists in the ‘Other’ facet, after which two new facets were established, Collaborative and Indiscernible, allowing all playlist names to be coded into at least one facet of the twelve facets. Following Thompson’s experience with Last.fm (2008, pp. 21-22), it was expected that user playlist names would deviate from standard industry names as they contained folksonomic naming conventions and so user playlist coding included two new facets, ‘Superlative’, and ‘Other Personal Meaning’.

**Analysis of playlists: qualitative and quantitative**

The playlists generated some challenges as some playlist names, particularly those categorised along distinct geographical, social, or cultural lines, possessed a meaning that could not be understood without additional social context. Coding therefore involved checking descriptions, dictionaries and generating some additional sub-categories to ensure sufficient meaning was captured. For example, within user playlists, some were simply named using emoji characters, such as “” or “,” though in these rare cases the Other Personal Meaning facet was able to accommodate these, though user meaning for list names is not possible to guarantee given cultural and contextual variations.

Data tables were generated from the raw coding data using the Pivot Tables feature in Microsoft Excel. Additional data was appended to the tables for analytical purposes, including a sum total of all codes applied to playlists, and a column showing the percentage of codes, by facet, as a percentage of the sum total of all codes.
Derivation of Genre data from the playlist data: generating a sub-sample

As genre data is not stored for playlists on Spotify beyond any indication provided in playlist categorisation or naming, genre data had to be derived from the artists of individual tracks within playlists. Due to the size of each dataset, this would be a prohibitively extensive undertaking if all tracks in all playlists were to be analysed. A suitable sampling of playlists thus needed to be determined. The decision was also taken to sample only editorial playlists based on data accessibility.

For editorial playlists, artist genre data was acquired using SpotiPy, a Python-based API wrapper, and the Generalized Spotify Analyser (GSA) to download data via Spotify’s Web API (Heggli, 2021). Within the GSA, the GSA_basicExample.py script was used to retrieve data on a playlist-by-playlist basis using the playlist URI as retrieved from the Spotify web player. Data was stored as a series of PKL files, which were first read as Pandas DataFrames, and then exported as CSV files using the IPython terminal in the IDE programme Spyder. Genres were provided for every track in a given playlist, determined on the basis of genres associated with the performing artist(s) in question.

The literature suggests that genre and context were two of the most important areas of focus for this study and sub-sampling focused on playlist contents as shown in Table 1 below:

Table 1. Overview of playlist sub-samples based on coding facet and number of playlists sampled per subsample.

Genre tags and unigrams taken from deconstructed tags were aggregated and analysed for each individual playlist. Unigrams also allowed for an additional dimension of analysis, helping to establish wider trends beyond the more granular tag-based classification and potentially highlight the prevalence of genre ‘families’ (e.g. rock, pop, rap, etc.) across many genre tags and to enable this extra processing split out genre tags and unigrams.

Playlist musicological property data and analysis

A series of musicological properties are stored on a track-by-track basis by Spotify and are retrievable via the Spotify API. This data was acquired concurrently with genre data. These ‘audio features’ comprise of several metrics (‘acousticness’, ‘valence’, ‘loudness’, etc.) and quantify...
musicological properties of given tracks. By measuring these features in numerical terms, a quantitative analysis of this data was possible.

The Analysis ToolPak in Microsoft Excel was used to perform a statistical analysis of audio feature data for each playlist. Of particular interest in this dataset was the standard deviation, which quantified the amount of spread per audio feature for each playlist. All playlist data was collated in a single spreadsheet along with tag and unigram frequency data, creating a convenient summary view allowing for overview of all data and highlighting of significant data points. Seven of the nine audio features use the same zero-to-one measurement scale; however, loudness and tempo are measured in decibels and beats per minute respectively.

Radar charts, created using RAWGraphs, provided a graphical comparison between levels of variation for different playlist sub-samples. The final data samples consisted of 5,180 playlists sampled for coding (including 4,180 editorial playlists, and 1,000 user playlists sampled from the Million Playlist Dataset). A sub-sample of 45 editorial playlists was selected for detailed quantitative analysis using data retrieved via the Spotify Web API resulting in 5,267 songs featuring 10,887 genre tags, with 1,458 songs tagged as ‘Unknown’ only. Disregarding ‘Unknown’ tags, 19,736 tag unigrams were used to describe songs across all playlists. A total of 1,136 unique genre tags were used to describe songs in the playlist sample. Finally, after these criteria forty-five editorial lists were selected as containing rich data on genre and that 45 was a manageable number for in-depth analysis.

Results

The playlists samples taken from Spotify were analysed in three ways: First, playlist names were coded to establish themes and determine prevalence of generic and non-generic names for editorial and user playlists. The second stream of analysis entailed looking at Spotify’s genre tag data for individual playlists within the sub-samples, focusing on the prevalence of the most-common tags and unigrams taken from tags applied to songs within those playlists, to determine levels of genre homogeneity per sub-sample. The final stream of analysis involved using Spotify’s ‘audio feature’
data to determine levels of musicological homogeneity within the sub-samples, with an aim to
determine whether non-genre based playlists shared a level of similarity beyond genre
classification.

**Playlist name data: Editorial playlists**

Table 4.1 shows the sum of codes by coding facet for all editorial playlists. A total of twelve facets
were created to classify codes. The column ‘Code Count’ reflects the total number of times codes of
a given facet were applied to a playlist name. The ‘Sum of Codes (%)’ column shows this value as a
percentage share of the total number of codes.

[INSERT TABLE 2 HERE]

*Table 2. Sum of codes for all editorial playlists*

The data shows generic codes applied to playlist names were the second-most prevalent of all
coding facets, accounting for 18.96% of codes across all editorial playlists. Context-based codes
were the most prevalent coding facet, accounting for 22.49% of codes.

Figure 1 shows instances of concurrent codes of playlist names per coding facet. This data shows
playlist names coded as generic were also coded as context-based a total of 507 times, this is the
highest level of concurrence across all coding facets.

Other high levels of concurrence with the genre facet included the ‘temporal quality’ facet (316),
the ‘popularity’ facet (213), the ‘musical property’ facet (195), and the ‘geographic origin’ facet
(159). These are the 5th, 7th, 8th and 10th highest levels of concurrence respectively, across all coding
facets within the editorial playlist sample.

[ INSERT FIGURE 1 HERE ]
Figure 1. Matrix chart showing distribution of concurrent codes of playlist names from the editorial playlist dataset for each coding facet. Darker green heatmap shading indicates higher frequencies of code concurrence. Cells shaded dark grey show the count of codes for a given facet, corresponding to the values seen in Table 2.

The total number of codes belonging to the ‘genre’ facet, combined with the total number of concurrent codes of playlist names belonging to all other facets totals 2,853 codes, equating to 35.58% of all codes within the sample. This highlights that while generic codes may not account for anywhere near a majority of all codes, genre is nonetheless a significant means of description for playlist names and is readily used in combination with other forms of description, not least the most common non-generic form of playlist description – context.

Playlist name data: User playlists

Table 3 shows the sum of codes by coding facet for the user playlist sample. This data includes the additional two coding facets created for playlist names relating to Superlative references and Other Personal meanings. Generic codes were the third-most prevalent of all coding facets, accounting for 12.08% of all codes across playlist names from the user playlist sample. It is notable that generic codes across the dataset was lower by 6.84% compared to the editorial playlist dataset. Context-based codes were most prevalent, accounting for 38.82% of all codes, followed by codes of the ‘Other Personal Meaning’ facet which accounted for 19.98% of all codes.

Table 3. Sum of codes for all user playlists.

Context-based codes account for a greater proportion of total codes in comparison to the editorial playlist dataset, and the gap between context-based and generic shares of codes is significantly larger. Context based codes account for almost double the proportion of the next-highest-ranked facet, codes ascribed to ‘Other Personal Meaning’, and three times the proportion of codes ascribed to the ‘Genre’ facet. Compared with the editorial playlist dataset, user playlists saw considerably less concurrent coding of playlist names. While 52.44% of all codes of editorial playlists were
applied in concurrence with another code for a given playlist, only 5.71% of codes of user playlists were applied in concurrence, suggesting an overall simplistic nature of user playlist names, with many names only drawing on one theme.

**Playlist name data: Genre tag data**

Table 4 shows the average prevalence of the most frequent genre tags and unigrams for each playlist sub-sample. ‘context only’ and ‘genre and context only’ sub-samples showed the highest levels of prevalence for both genre tags and unigrams.

The ‘genre and context only’ sub-sample showed the outright highest prevalence for average tag count (31.84%) and unigram count (35.66%) across its playlists, closely followed by the ‘context only’ sub-sample (31.61% and 31.67%, respectively). While the ‘genre only’ subsample showed significantly lower levels of prevalence (18.50% and 25.49% respectively), it did show the highest unigram count per track at 1.77.

| INSERT TABLE 4 HERE |

**Table 4. Average prevalence of most frequent genre tags and unigrams for playlist sub-samples.**

Tables 5, and 6 show the top ten highest instances of tag counts and unigram counts for individual playlists, respectively.

| INSERT TABLE 5 HERE |

**Table 5** Top ten highest ranking playlists for frequency of most frequently occurring genre tag.

| INSERT TABLE 6 HERE |

**Table 6** Top ten highest ranking playlists for frequency of most frequently occurring unigram taken from genre tags.

When observing average tag count per track, the higher tag count per track for the ‘genre only’ sub-sample in comparison to other sub-samples is clearly demonstrated. Meanwhile, the ‘abstract
quality only’ sub-sample evidently lacked genre tag homogeneity. The ‘abstract quality only’ sub-sample had significantly lower frequencies of homogeneity in genre tags, and to a lesser extent unigrams, showing average frequencies of 10.16% and 15.71% for the most common tags and unigrams in each playlist. This sub-sample included five playlists whose names were determined to possess a quality separate from musicological properties of the playlist, addressing an object, concept or a non-sound based quality. Table 7 shows overall rankings for genre tag and unigram frequency for each of these playlists.

Table 7 Overall rankings of most-common genre tag and unigram frequency for playlists from the ‘abstract quality ONLY’ sub-sample.

Playlists from all samples were ranked, out of a total of 45 playlists, from highest to lowest frequencies for each metric. The table shows that playlists from the sub-sample rank particularly poorly for the frequency of their most-common genre tags, with only one playlist ranking outside of the bottom ten overall, while only two playlists rank outside of the bottom ten for frequency of their most-common unigrams. This implies consistently low levels of genre homogeneity for the sub-sample. Overall, the sub-samples showed similar differences between overall average counts and average counts per track for unigrams as were seen for tags, though the rates for each increased for the ‘genre only’ sub-sample to the extent that it was able to rank highest for average unigram count per track across all sub-samples.

Playlist name data: Musicological property data

Playlists of the ‘genre and context only’ sub-sample showed the lowest average standard deviations for four of the nine metrics (Danceability, Energy, Acousticness, and Valence) while the ‘geographic region only’ and ‘musical quality only’ sub-samples showed the lowest standard deviations for two metrics each. Playlists of the ‘context only’ sub-sample showed the lowest average standard deviation for Tempo.
Table 8. Average standard deviations of Spotify’s audio features across playlists per sub-sample. Lowest standard deviations across sub-samples for each metric are shaded green.

The data in Table 8 was normalised and offset for visualisation radar chats in Figure 2. In these charts, values closer to zero indicate the smallest deviations from the mean for playlists within these sub-samples. This visualisation displays lower levels of variation in context-based sub-samples. While all sub-samples show low standard deviations for at least one feature, context-based sub-samples exhibit several, with the ‘genre and context only’ sub-sample only shows a larger standard deviation for Tempo, and the ‘context only’ subsample having a high standard deviation for Instrumentalness and a middling one for Valence.

[INSERT FIGURE 2 HERE]

Figure 2. Radar charts showing normalised and offset standard deviations of audio feature data for playlist sub-samples.

This data suggests context-based playlists are more likely to exhibit consistencies in musicological properties than other forms of playlist classification. Speaking specifically, playlists of the ‘genre and context only’ sub-sample showed little variation in the Danceability, Energy, Speechiness, Acousticness and Valence audio features. Averages of normalised standard deviations show context-based playlists demonstrating the most consistency in their audio properties when ranked amongst other sub-samples. The ‘genre and context only’ subsample showed the lowest average standard deviation of 0.500, while the ‘abstract quality only’ sub-sample showed the highest, at 1.852. While the ‘genre and context only’ sub-sample showed the outright lowest standard deviations for four audio features, and within the ‘context only’ subsample are lowest ranked for eight of nine audio features across all playlists sampled for the study, as Table 9 shows.

[INSERT TABLE 9 HERE]

Table 9. Ranking of standard deviation of individual playlists from the ‘context ONLY’ sub-sample amongst all playlists.
Figure 3 shows levels of variance for audio feature data for the ‘Night Rider’ playlist and provides an example of one editorial playlist in the analysis with a relatively low level of homogeneity in genre tags and unigrams. Overall, the playlist has the lowest level of variation for tempo across all sub-samples, and low levels of variation in the speechiness, instrumentalness, danceability, and loudness audio features, showing consistency in musicological properties can exist in the absence of either generic playlist naming or genre tag homogeneity.

[INSERT FIGURE 3 HERE]

**Figure 3. Radar chart showing normalised and offset standard deviations for ‘audio feature’ data for the ‘Night Rider’ editorial playlist. Values closer to zero represent lowest levels of variation for these metrics across songs constituting the playlist.**

**Discussion**

Coding of editorial and user playlists on Spotify revealed a breadth of descriptive themes. Generic terms formed a sizeable minority for editorial and user playlist samples, though it would also be valid to shift perspective and state that non-genre themes were applied in playlist naming in 81.04% and 87.92% of cases, respectively. Pachet and Cazaly (2002) argued that increasingly diffuse terminology employed by the music industry was a means to create the most straightforward route to music for the consumer. This is evident with Spotify’s packaging and presentation of playlists. Colourful, single-word categories play on emotions and situations as well as musical themes.

While surface-level data depicts a relatively minor role for genre in playlist music classification, deeper analysis reveals nuance. Genre accounted for the second greatest number of codes for the editorial playlist dataset, and the third greatest for the user playlist dataset. Furthermore, concurrence of themes was a notable feature of playlist names involving genre, especially for editorial playlists. Regev’s (2002) remarks on the ‘pop-rockization’ of genre terms casting a wide net encompassing many themes resonate with this data, and further tie into Negus’ (1999) discussion of the boundary-breaking properties of certain genre forms, suggesting genre remains a powerful descriptive tool for music collections despite its overall importance in music classification waning.
The user playlist dataset shows genre as having a slimmer share of codes. This could be explained by the rise in context-based listening as music and listening technologies become increasingly pervasive (Morris (2015); Huber (2019), and the use of playlists to replicate rituals of physical collecting (Hagen, 2015), representing a departure from traditional generic forms of classification.

In both datasets concurrent codes were analysed, with Editorial playlists far more likely to show a concurrence of themes within playlist names. User playlist names were much more straightforward, typically drawing on a single theme in their naming. While playlist-based classification may have enabled users to employ more diverse themes when naming playlists, these will typically be focused on a single event, memory, or purpose. Editorial playlists, by comparison, are required to appeal to vast audiences. Eriksson, et al. (2019) detailed the importance of playlists for Spotify in terms of advertising-based revenue generation. It would stand to reason that MSPs would aim to describe playlists using broad strokes to widen their appeal.

Several user playlist names were classified in the ‘Other Personal Meaning’ facet. 19.98% of all playlist names were coded as featuring this theme, making this the second-highest ranked facet, below context (38.82%) but above genre (12.08%). This echoes Hagen (2015) who found that users applied personalised names to their playlists which, without any knowledge of the context of naming, are impossible to code in a more specific way. Individual users were far more likely to apply highly personalised names (e.g. ‘Flowers’, ‘Michael’, ‘<3 <3 <3’ and ‘Strawberry Jam’) to playlists in comparison to the ‘catch-all’ appeal of editorial playlists.

A surface analysis of playlist names paints a picture of a diverse collection of non-generic themes being used to describe playlists, yet analysis reveals that that genre themes in playlist naming were not necessarily the most effective predictor of genre homogeneity within playlists. Frith’s (1996), discussion of the music industry’s use of genre as a means of bridging the gap between product offerings and consumer demand, appear pertinent here. Frith detailed the creation of abstract categories in record stores to market music, and even the creation of entire genres such as ‘world music’ in the pursuit of the ‘fantasy consumer’, and this perhaps intensifies with narrowcasting and technology changes (see also: Taylor, 1997, Kusek and Leonhard, 2005; Morris, 2015; Rossman, 2012). What appears to be happening in the streaming context is the packaging of generically homogenous music, using non-generic terminology that appeals to the needs of the modern consumer. Just as Frith describes the importance of effective categorisation in the context of the record store in decades past, Spotify now faces identical challenges as it seeks to exploit the
important playlist format (Aguiar and Waldfogel, 2018; Prey et al., 2022). With this in mind, to speak of a decline in genre’s relevance, as suggested by the evidence in the previous section, would neglect the presence of genre homogeneity demonstrated by high frequencies of similar genre tags and unigrams in playlists that do not bear generic themes in their naming.

The ‘abstract quality only’ sub-sample suggested lower frequency of homogeneity in genre tags and a similar, but lesser, finding with unigrams which is curiously out of step with other sub-samples. The sub-sample’s musicological properties may partially explain this as five playlists were identified as being named for their sound-based quality rather than other features. Exploring this phenomenon may reveal more about the apparently strong relationship between genre and context. Zhang and Olson (2015) discuss the notion of genre possessing an essence-context duality. While they suggest context is a property innate to genre, it could be posited that context, when applied as an extrinsic descriptor of playlists in combination with genre terms, can in fact strengthen levels of genre homogeneity to levels beyond those seen in playlists with purely generic names.

Playlist names of the ‘abstract quality only’ sub-sample lack reference to either genre, context, or seemingly any reference to music whatsoever. They lack the properties which would help to increase levels of genre homogeneity in the playlist. By contrast, playlist names based on genre and/or context themes show significantly higher likelihood of presenting genre homogeneity.

A conclusion can be drawn that non-generic playlist naming is not necessarily an indicator of a lack of genre homogeneity. The presence of context-based themes in playlist naming appears to be the strongest predictor of genre homogeneity, suggesting there is still a role to play for genre in playlist-based music classification, even if genre terms are not directly employed within playlist names.

The objective in analysing musicological property data was to determine presence of trends in musicological properties within playlists. This was done using the same six sub-samples as those used when studying genre tags. Much discussion on the treatment of popular music from an information science perspective focuses on neglect of the cultural aspect in classification due to supposed historical biases towards traditional music forms. The result is classification schemes focusing primarily on the musicological properties and instrumentation of music, rather than any cultural aspect. However, such commentaries (Hemmasi and Young (2000); Abrahamsen (2003), raise the question of whether playlist-based forms of music classification exhibit any consistencies in musicological properties beyond purely generic descriptions, essentially creating ‘genre-by-
proxy’, playlists containing songs sharing distinct musicological properties while not necessarily sharing generic similarities.

Analysis of ‘audio feature’ data revealed the ‘genre and context only’ sub-sample showed the lowest levels of variation in musicological properties. Standard deviations of ‘audio feature’ measurements showed this sub-sample had the most consistency in four of the nine metrics used: Danceability, Energy, Acousticness, and Valence (Table 8). The comparison of audio features (normalised and offset) showed that overall, the ‘genre and context only’ sub-sample had the lowest average (0.500). This was followed by the ‘context only’ sub-sample (0.833), the ‘genre only’ sub-sample (1.187), ‘geographic region only’ (1.327), ‘musical quality only’ (1.596), and finally ‘abstract quality only’ (1.852). The radar charts shown in Figure 4 give the clearest visualisation of this data for each ‘audio feature’.

While genre may be absent as a theme in playlist naming, there are strong indications based on context-based playlists that musicological properties can exist despite this absence. Siles et al. (2019) investigated the notion of playlists existing as ‘affective genres’, and argued that by utilising the power of affect, new roles can be discovered for music exceeding the limitations of traditional genre. While Siles et al. focused on user-based playlist usage, the evidence from the present study would suggest editorial playlists can also be utilised by MSPs to harness the potential to cultivate affect, creating playlists whose names tap into moods and activities, but whose contents serve as a means of amplifying feeling to create what could be regarded as an affective genre separate from traditional notions associated with the term.

**Conclusion**

Are playlists, as De Regt (2020) described, ‘the new genre’? This study has scrutinised this claim, first by framing genre’s evolving role in music classification, and developing an understanding of how playlist-based classification and novel listening technologies have helped to change the ways MSPs and their users engage with their music collections. The limitations of the study as exploratory in nature, the lack of automatic genre and audio feature data harvesting using the Spotify API due to technical difficulties, and in terms of the sample size a sample of 5180 subsampled down to 45 limit the studies generalisability. However, the study suggests genre can still play a significant role in playlist-based classification, but it is a single tree in a dense forest of themes used by MSPs and users to describe their collections. However, even playlists bearing non-generic descriptions,
particularly context-based themes, often possess a high degree of genre homogeneity, suggesting a form of genre classification by another name. The Study of musicological properties showed context-based playlists were also likely to show high levels of similarity in these properties both with and without genre homogeneity, resonating with user-based studies that have discovered evidence of playlists being used to create ‘affective’ genres as a means to amplify mood. In multiple ways this study has helped to qualify and quantify the diverse ways in which the music industry and users interact with playlists as a novel form of music classification.

References


De Regt, L. (2020). *Spotify Playlists Are The New Genre*, available at:


Forde, E. (2017). 'They Could Destroy the Album': How Spotify's Playlists have Changed Music for Ever, available at:


Full Page. (2022). *GoFullPage*, available at:


<table>
<thead>
<tr>
<th>Playlist by coding facet</th>
<th>No. of playlists sampled for detailed analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract Quality ONLY</td>
<td>5</td>
</tr>
<tr>
<td>Context ONLY</td>
<td>10</td>
</tr>
<tr>
<td>Genre AND Context ONLY</td>
<td>10</td>
</tr>
<tr>
<td>Genre ONLY</td>
<td>10</td>
</tr>
<tr>
<td>Geographic Origin ONLY</td>
<td>5</td>
</tr>
<tr>
<td>Musical Property ONLY</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total playlists in sub-sample</strong></td>
<td><strong>45</strong></td>
</tr>
</tbody>
</table>

*Table 1. Overview of playlist sub-samples based on coding facet and number of playlists sampled per subsample.*
### Table 2

<table>
<thead>
<tr>
<th>Values</th>
<th>Code Count</th>
<th>Sum of Codes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of Abstract Quality</td>
<td>189</td>
<td>2.55%</td>
</tr>
<tr>
<td>Sum of Branding / Promo</td>
<td>488</td>
<td>6.60%</td>
</tr>
<tr>
<td>Sum of Collaborative</td>
<td>319</td>
<td>4.31%</td>
</tr>
<tr>
<td>Sum of Context</td>
<td>1663</td>
<td>22.48%</td>
</tr>
<tr>
<td>Sum of Genre</td>
<td>1402</td>
<td>18.95%</td>
</tr>
<tr>
<td>Sum of Geographic Origin</td>
<td>584</td>
<td>7.89%</td>
</tr>
<tr>
<td>Sum of Indiscernible</td>
<td>4</td>
<td>0.05%</td>
</tr>
<tr>
<td>Sum of Musical Property</td>
<td>702</td>
<td>9.49%</td>
</tr>
<tr>
<td>Sum of Named Artists</td>
<td>691</td>
<td>9.34%</td>
</tr>
<tr>
<td>Sum of Non-Musical Sound</td>
<td>65</td>
<td>0.88%</td>
</tr>
<tr>
<td>Sum of Popularity</td>
<td>658</td>
<td>8.89%</td>
</tr>
<tr>
<td>Sum of Temporal Quality</td>
<td>634</td>
<td>8.57%</td>
</tr>
<tr>
<td><strong>Sum of all codes</strong></td>
<td><strong>7399</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Sum of codes for all editorial playlists
Figure 1. Matrix chart showing distribution of concurrent codes of playlist names from the editorial playlist dataset for each coding facet. Darker green heatmap shading indicates higher frequencies of code concurrence. Cells shaded dark grey show the count of codes for a given facet, corresponding to the values seen in Table 2.
TABLE 3

<table>
<thead>
<tr>
<th>Values</th>
<th>Code Count</th>
<th>Sum of Codes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of Abstract Quality</td>
<td>17</td>
<td>1.62%</td>
</tr>
<tr>
<td>Sum of Branding / Promo</td>
<td>7</td>
<td>0.67%</td>
</tr>
<tr>
<td>Sum of Collaborative</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Sum of Context</td>
<td>408</td>
<td>38.82%</td>
</tr>
<tr>
<td>Sum of Genre</td>
<td>127</td>
<td>12.08%</td>
</tr>
<tr>
<td>Sum of Geographic Origin</td>
<td>10</td>
<td>0.95%</td>
</tr>
<tr>
<td>Sum of Indiscernible</td>
<td>2</td>
<td>0.19%</td>
</tr>
<tr>
<td>Sum of Musical Property</td>
<td>41</td>
<td>3.90%</td>
</tr>
<tr>
<td>Sum of Named Artists</td>
<td>41</td>
<td>3.90%</td>
</tr>
<tr>
<td>Sum of Non-Musical Sound</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Sum of Other Personal Meaning</td>
<td>210</td>
<td>19.98%</td>
</tr>
<tr>
<td>Sum of Popularity</td>
<td>16</td>
<td>1.52%</td>
</tr>
<tr>
<td>Sum of Superlative</td>
<td>63</td>
<td>5.99%</td>
</tr>
<tr>
<td>Sum of Temporal Quality</td>
<td>109</td>
<td>10.37%</td>
</tr>
<tr>
<td><strong>Sum of all codes</strong></td>
<td><strong>1051</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Table 3. Sum of codes for all user playlists.*
<table>
<thead>
<tr>
<th>Playlist Sub-sample</th>
<th>Tag Frequency</th>
<th>Unigram Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg. Tag Count (%)</td>
<td>Avg. Count Per Track</td>
</tr>
<tr>
<td>abstract quality ONLY</td>
<td>10.16%</td>
<td>0.28</td>
</tr>
<tr>
<td>context ONLY</td>
<td>31.61%</td>
<td>0.42</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>31.84%</td>
<td>0.62</td>
</tr>
<tr>
<td>genre ONLY</td>
<td>18.50%</td>
<td>0.54</td>
</tr>
<tr>
<td>geographic origin ONLY</td>
<td>18.72%</td>
<td>0.39</td>
</tr>
<tr>
<td>musical quality ONLY</td>
<td>17.72%</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Table 4. Average prevalence of most frequent genre tags and unigrams for playlist sub-samples.
TABLE 5

<table>
<thead>
<tr>
<th>Playlist Sub-sample</th>
<th>Playlist Name</th>
<th>Most Freq. Tag</th>
<th>Tag Count (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>context ONLY</td>
<td>Sleep</td>
<td>sleep</td>
<td>86.00%</td>
</tr>
<tr>
<td>context ONLY</td>
<td>Reading Soundtrack</td>
<td>background piano</td>
<td>80.81%</td>
</tr>
<tr>
<td>context ONLY</td>
<td>Keep Calm</td>
<td>background music</td>
<td>51.08%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>Workday Jazz</td>
<td>background jazz</td>
<td>76.25%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>WORKOUT</td>
<td>k-pop</td>
<td>48.39%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>Beast Mode Christian</td>
<td>christian hip hop</td>
<td>35.84%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>Rainy Day Jazz</td>
<td>background jazz</td>
<td>35.58%</td>
</tr>
<tr>
<td>genre ONLY</td>
<td>Future Funk</td>
<td>future funk</td>
<td>67.74%</td>
</tr>
<tr>
<td>geographic origin ONLY</td>
<td>Made in Kenya</td>
<td>kenyanpop</td>
<td>32.50%</td>
</tr>
<tr>
<td>musical quality ONLY</td>
<td>Soft Piano</td>
<td>neo-classical</td>
<td>30.51%</td>
</tr>
</tbody>
</table>

Table 5 Top ten highest ranking playlists for frequency of most frequently occurring genre tag.
<table>
<thead>
<tr>
<th>Playlist Sub-sample</th>
<th>Playlist Name</th>
<th>Most Freq. Unigram</th>
<th>Unigram Count (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>context ONLY</td>
<td>Sleep</td>
<td>sleep</td>
<td>77.48%</td>
</tr>
<tr>
<td>context ONLY</td>
<td>Reading Soundtrack</td>
<td>background</td>
<td>44.12%</td>
</tr>
<tr>
<td>context ONLY</td>
<td>Work From Home</td>
<td>pop</td>
<td>40.99%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>Country Cookout</td>
<td>country</td>
<td>49.22%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>WORKOUT</td>
<td>k-pop</td>
<td>48.16%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>Workday Jazz</td>
<td>jazz</td>
<td>47.83%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>Rainy Day Jazz</td>
<td>jazz</td>
<td>42.71%</td>
</tr>
<tr>
<td>genre AND context ONLY</td>
<td>Salsa Romantica</td>
<td>salsa</td>
<td>36.91%</td>
</tr>
<tr>
<td>genre ONLY</td>
<td>Boogie Rock</td>
<td>rock</td>
<td>39.51%</td>
</tr>
<tr>
<td>genre ONLY</td>
<td>Future Funk</td>
<td>future</td>
<td>39.15%</td>
</tr>
</tbody>
</table>

Table 6 Top ten highest ranking playlists for frequency of most frequently occurring unigram taken from genre tags.
TABLE 7

<table>
<thead>
<tr>
<th>Playlist Name</th>
<th>Most Freq. Tag</th>
<th>Tag Count (%)</th>
<th>Rank</th>
<th>Most Freq. Unigram</th>
<th>Unigram Count (%)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark &amp; Stormy</td>
<td>indie folk</td>
<td>10.14%</td>
<td>35</td>
<td>pop</td>
<td>14.99%</td>
<td>38</td>
</tr>
<tr>
<td>Glimmer</td>
<td>electropop</td>
<td>8.43%</td>
<td>39</td>
<td>pop</td>
<td>15.77%</td>
<td>37</td>
</tr>
<tr>
<td>Halo</td>
<td>electronica</td>
<td>7.69%</td>
<td>42</td>
<td>house</td>
<td>8.81%</td>
<td>45</td>
</tr>
<tr>
<td>Lush + Ethereal</td>
<td>indie folk</td>
<td>18.73%</td>
<td>18</td>
<td>indie</td>
<td>19.72%</td>
<td>28</td>
</tr>
<tr>
<td>Space</td>
<td>rock</td>
<td>5.83%</td>
<td>44</td>
<td>rock</td>
<td>19.26%</td>
<td>30</td>
</tr>
</tbody>
</table>

*Table 7* Overall rankings of most-common genre tag and unigram frequency for playlists from the 'abstract quality ONLY' sub-sample.
<table>
<thead>
<tr>
<th>Audio Feature (Standard Deviation)</th>
<th>Playlist Sub-Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Erre ONLY</td>
</tr>
<tr>
<td>Acousticness</td>
<td>0.175</td>
</tr>
<tr>
<td>Danceability</td>
<td>0.111</td>
</tr>
<tr>
<td>Energy</td>
<td>0.134</td>
</tr>
<tr>
<td>Instrumentalness</td>
<td>0.233</td>
</tr>
<tr>
<td>Liveness</td>
<td>0.151</td>
</tr>
<tr>
<td>Loudness (dB)</td>
<td>2.553</td>
</tr>
<tr>
<td>Speechiness</td>
<td>0.074</td>
</tr>
<tr>
<td>Tempo (BPM)</td>
<td>23.408</td>
</tr>
<tr>
<td>Valence</td>
<td>0.182</td>
</tr>
</tbody>
</table>

Table 8. Average standard deviations of Spotify’s audio features across playlists per sub-sample. Lowest standard deviations across sub-samples for each metric are shaded green.
FIGURE 2

Figure 2. Radar charts showing normalised and offset standard deviations of audio feature data for playlist sub-samples.
TABLE 9

<table>
<thead>
<tr>
<th>Audio Feature</th>
<th>Lowest-Ranked Playlist</th>
<th>Overall Rank</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acousticness</td>
<td>Reading Soundtrack</td>
<td>1</td>
<td>0.010</td>
</tr>
<tr>
<td>Danceability</td>
<td>Sleep</td>
<td>1</td>
<td>0.055</td>
</tr>
<tr>
<td>Energy</td>
<td>Sleep</td>
<td>1</td>
<td>0.019</td>
</tr>
<tr>
<td>Instrumentalness</td>
<td>Mood Booster</td>
<td>5</td>
<td>0.003</td>
</tr>
<tr>
<td>Liveness</td>
<td>Reading Soundtrack</td>
<td>1</td>
<td>0.014</td>
</tr>
<tr>
<td>Loudness (dB)</td>
<td>Afterhours</td>
<td>1</td>
<td>0.304</td>
</tr>
<tr>
<td>Speechiness</td>
<td>Sleep</td>
<td>1</td>
<td>0.005</td>
</tr>
<tr>
<td>Tempo (BPM)</td>
<td>Night Rider</td>
<td>1</td>
<td>2.202</td>
</tr>
<tr>
<td>Valence</td>
<td>Sleep</td>
<td>1</td>
<td>0.045</td>
</tr>
</tbody>
</table>

Table 9. Ranking of standard deviation of individual playlists from the ‘context ONLY’ sub-sample amongst all playlists.
Figure 3. Radar chart showing normalised and offset standard deviations for ‘audio feature’ data for the ‘Night Rider’ editorial playlist. Values closer to zero represent lowest levels of variation for these metrics across songs constituting the playlist.