Parsing Consumption Preferences of Music Streaming Audiences

through Concatenating Data Analytics

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie der Ludwig-Maximilians-Universität München

vorgelegt von

Sophie Brüggemann

aus

Augsburg

2020

Referent/in: Prof. Dr. med. Dr. phil. Lorenz Welker

Korreferent/in: Prof. Dr. phil. Irene Holzer

Tag der mündlichen Prüfung: 15.05.2020

Parsing Consumption Preferences of Music Streaming Audiences

through Concatenating Data Analytics

Sophie Brüggemann

Department of Musicology

Ludwig-Maximilians-University

This dissertation is submitted for the degree of

Doctor of Philosophy

Munich January 2020

To every relationship in science, as in humanity, the same principle applies: never forget to bring along a box full of time and attention.

I dedicate this thesis to my loving parents and grandparents, who have supported me throughout the years and encouraged my insatiable curiosity in the arts and in science.
Thank you for always sparing some time and attention for me and my theories.

Acknowledgements

I would like to acknowledge my supervisor Professor Dr. med. Dr. phil. Lorenz Welker for inspiring my initial interest in music psychology and especially music listening behaviors. Furthermore, I would like to thank Dr. phil. Magdalena Zorn for having an open ear and setting an example as an academic role model.

Finally, my work experience at Nielsen Music and Spinnin' Records was instrumental in defining the path of my research and both companies supported me as a cooperation partner.

Abstract

As demands for insights on music streaming listeners continue to grow, scientists and industry analysts face the challenge to comprehend a mutated consumption behavior, which demands a renewed approach to listener typologies. This study aims to determine how audience segmentation can be performed in a time-relevant and replicable manner. Thus, it interrogates which parameters best serve as indicators of preferences to ultimately assist in delimiting listener segments.

Accordingly, the primary objective of this research is to develop a revised typology that classifies music streaming listeners in the light of the progressive phenomenology of music listening. The hypothesis assumes that this could be solved by positioning listeners – rather than products – at the center of streaming analysis and supplementing sales- with user-centered metrics. The empirical research of this paper was based on grounded theories, enriched by analytical case studies. For this purpose, behavioral and psychological research results were interconnected with market analysis and streaming platform usage data.

Analysis of the results demonstrates that a concatenation of multi-dimensional data streams facilitates the derivation of a typology that is applicable to varying audience pools. The findings indicate that for the delimitation of listener types, the motivation, and listening context are essential key constituents. Since these variables demand insights that reach beyond existing metrics, descriptive data points relating to the listening process are subjoined. Ultimately, parameter indexation results in listener profiles that offer novel access points for investigations, which make imperceptible, interdisciplinary correlations tangible. The framework of the typology can be consulted in analytical and creational processes. In this respect, the results of the derived analytical approach contribute to better determine and ultimately satisfy listener preferences.

Zusammenfassung

Während die Nachfrage nach Erkenntnissen über Musik-Streaming-Hörer kontinuierlich steigt, stehen Wissenschaftler sowie Industrieanalysten einem geänderten Konsumptionsverhalten gegenüber, das eine überarbeitete Hörertypologie fordert. Die vorliegende Studie erörtert, wie eine Hörersegmentierung auf zeitgemäße und replizierbare Weise umgesetzt werden kann. Demnach beschäftigt sie sich mit der Frage, welche Parameter am besten als Indikatoren für Hörerpräferenzen dienen und wie diese zur Abgrenzung der Publikumssegmente beitragen können.

Dementsprechend ist es das primäre Ziel dieser Forschung, eine überarbeitete Typologie aufzustellen, die Musik-Streaming-Hörer in Anbetracht der progressiven Erscheinungsform des Musikhörens klassifiziert. Die Hypothese nimmt an, dass dies realisierbar ist, wenn der Hörer – anstelle von Produkten – im Zentrum der Streaming-Analyse steht und absatzzentrierte durch hörerzentrierte Messungen ergänzt werden. Die empirische Forschung basiert auf systematischen Theorien, untermauert durch analytische Fallbeispiele. Hierfür werden psychologische und verhaltenswissenschaftliche Forschungserkenntnisse mit Marktanalysen und Nutzerdaten von Musikstreaming-Portalen fusioniert.

Die Analyse der Ergebnisse verdeutlicht, dass eine Verkettung von multidimensionalen Rohdaten die Erhebung einer Typologie ermöglicht, die auf mehrere Hörergruppen anwendbar ist. Die Befunde signalisieren, dass die Hörmotivation und der Hörkontext bei der Abgrenzung der Publikumstypen Schlüsselelemente darstellen. Da diese Variablen spezifische Kenntnisse fordern, die über vorliegende Kennzahlen hinausgehen, werden deskriptive Datenpunkte über den Hörvorgang ergänzt. Letztlich, resultiert die Indexierung der Parameter in Hörerprofilen, die neue Zugangspunkte für Untersuchungen bieten, die nicht ersichtliche, interdisziplinäre Korrelationen greifbar machen. Das Gerüst der Hörertypologie kann sowohl in Erstellungs- als auch in Analyseprozessen herangezogen werden. Somit tragen die Ergebnisse der entwickelten Analysemethode zum Verständnis und letztlich zur Erfüllung von Hörerpräferenzen bei.

Table of Contents

List of Figures

List of Tables

Nomenclature

Mathematical Measures

IQR	Interquartile range
POD	Power of differentiation
RII	Relative importance index

Technical Terms

API	Application program interface
GMBI	General music branding inventory
MIR	Music information retrieval

Other Acronyms

IFPI	International Federation of the Phonographic Industry

Chapter 1

Introduction

In 2017, music streaming overtook physical sales for the first time, accompanied by significant changes in the way music was consumed (IFPI, 2018). This transformation of music listening is revealed in the form of access-driven content configurations, bi-polar interaction potentials, and contextual perceptions in accordance with audience preferences. Based on their highly individual nature, those changes demand new listener-centric indicators of performance. To observe this, marketing and product managers, as well as analysts and market researchers, are increasingly striving to understand listeners beyond mere sales numbers. However, current analytical tools primarily deal with product- and sales-centric metrics. The absence of typologies that are contemporary, pertinent and applicable to varying audience pools has been caused by the lack of scientific research on changes in music consumption, listening behavior, and perception following the rise of music streaming services. Due to this knowledge gap, industry analytics typically establish audience segmentations based on outliers or clusters from sales data. The here established research aims to reverse this approach, starting with human-centric factors and substantiating them with data afterward.

The reformation of current analytical models could be accomplished by turning to listener-based audience partitions as utilized in media science, music sociology, and music psychology. However, those methodologies need to be revised to be applicable to a context-dominated digital music streaming experience. Thus, reassessing the conditions surrounding the act of listening is an initial step toward answering the overarching question of which parameters best serve as indicators to identify listener categories. Furthermore, statistical analyses of music streaming data and the incorporation of correlation theories can substantiate audience segments. In essence, the hypothesis of this paper presumes that the establishment of a newly revised audience typology, including consumer data points, paired with a context-centric typology, can enhance understanding of audience demands and ultimately contribute to audience satisfaction.

While various factors affect the intent of users and ways in which they listen to music, this study focuses on the perspectives of socio-psychology and media science. While some elements of the research question have been examined in discipline-specific papers, they have yet to be connected. Overall, inter-disciplinary approaches have only been implemented by few researchers, as for instance by P. Farnsworth, who aimed for this in the field of music psychology (Farnsworth, 1958). He expressed interest in studying music not only for its form and function, but also concerning the context of performance and its effects on the audience. He did so by undertaking systematic investigations through a quantitative research model, including statistical techniques, to generalize observations across individuals. However, this approach could not be maintained due to the rejection of experimental studies based on critique of non-replicable testing conditions and controversial methodologies. Consequently, research turned to qualitative techniques, such as semi-structured interviews, participant observation, and the use of musical examples (Clarke and Cook, 2004). In general, studies of music behavior combining scientific and applied analytics are limited, which calls for lateral thinking in empirical research. Thereby musical preference, as well as listener typology, can assist as common denominators, since they already serve as important empirical instruments for investigations across music sociology and psychology.

Existing methodologies present some difficulties that the following analysis aims to resolve. The goal is to first increase the volume, consistency, and relevance of samples by deriving them through a non-disturbing data acquisition process, thus preserving a context close to reality. Second, this new approach aims to make qualitative correlates visible within a quantitative process. Third, it aspires to move away from demographic criteria of differentiation by adding superordinate categories, such as motivation and context. Those account for a more flexible, behavior-based framework suitable to encompass contemporary listening habits. Analyzing datasets from a music streaming database with a newly derived multi-dimensional analysis that allows substantiation of this typology. In this way, the advanced methodology aims to contribute to closing the gap in the existing literature. As well as providing enhanced and up-to-date methodologies, this thesis sets itself apart from previous research with two objectives that aim to induce a broad shift in the measurement and assessment of digital music consumption. The first objective is to establish context and motivation as the key constituents of an audience typology. The second objective is to demonstrate the value of supplementing prevalent sales-centered consumption analysis in the field of digital music with a listener-centered approach.

For analysis, as typical for investigations in systematic musicology, the research touches upon several adjacent disciplines. Those are foremost psychology, sociology, cultural and media sciences, as well as economics via industry reports. The results carry forward research in the field of socio-psychology (Juslin, 2013; Kassabian, 2013; Sloboda, 2012), media science (Miles, 2018; Prey, 2017), and analytical industry reports (IFPI, 2018), all of which take diverse approaches, while so far serving a delimited set of subject-specific purposes.

As for the research background, media psychology has long concentrated on theories of media effects in television (Winterhoff-Spurk, 1999). R. Mangold et al. (2004) made first attempts to expand and systematize the area of research, although music remained largely a side factor. Consequentially, music psychology has for long mostly been found in research context concerning communicators, recipients, and feelings. But existing research of those areas, including Barrett (2006), Cespedes-Guevara and Eerola (2018), Juslin (2013), Nawaz et al. (2018), Pessoa (2008), Brosch et al. (2010), can help to connect music consumption in old and new media contexts from a psychological standpoint. In pairing such insights with product configurations and usage trends, the music consumption of the streaming age can be explored. The three-part division of music streaming aspects into consumption, behavior and psychology, allow to dissect how music is currently listened to and received. Those are underpinned by application-driven examples which are based on recent studies on the modern consumer culture in music (Miles, 2018; Nylund-Hagen, 2016; Prey, 2017; Smudits, 2007), new product configurations and distribution models (Wikström, 2012), as well as music specific studies on interaction and reception (Kassabian, 2013; Sloboda, 2012).

Alongside the consumer culture, the sociology of music has likewise developed in many directions in the past century, with T. W. Adorno as its most important representative. Even though, his normative orientation often became the object of criticism, the approaches he raised are still effective today. Especially in the 1990s and 80s the reception of his works increased, to a big extend due to his connection of music and social structures, awareness and cognition. T. W. Adorno's work is pervaded by the social function he attributes to music, which is marked by his criticism of the culture and music industry in *Dialektik der Aufklärung* (Horkheimer and Adorno, 1981). Accordingly, the concept of music under discussion has to be seen in context of a long tradition of criticism of the culture and music industry. This research deals with a broad concept of music, encompassing all music genres, whereby factors of mass-consumption on streaming platforms form central elements. Listening is thereby a multi-sensory experience, as reciprocal effects from multiple senses can be involved. Furthermore, it is assumed that music-intrinsic and non-musical aspects exert equal influence on listener preferences. Its variations are determined by the specific needs of each listener type, summarized by prototypical listening contexts and intents.

For a long time, the description of music scenes and their audiences has been based on studies conducted by radio stations and market research institutes with the aim of discussing listening preferences (Müller et al., 2002). However, new configurations of musical products require a new methodology for preference surveying. Towards the end of the 19th century, at the beginning of musical preference research, the research subject was different from todays: Concert goers rated music in questionnaires (Gembris, 1999), the frequency of performance of works or composers was recorded (Mark, 1998), or audio samples were evaluated using adjective scales (Brömse and Kötter, 1971). Nevertheless, those studies made it possible to collect the first data-based listener typologies on the basis of preference bundles and also clusters (Behne, 1986). Due to the brevity of music examples, however, this only allowed for the examination of already established music preferences and resulted in musical products where widely popular pieces appeared repetitively (Münch, 1998). North and Hargreaves (1997) reacted to this by pairing the evaluation of listening preferences with the correlation of familiarity and popularity levels, which are also included in the following analysis.

The listener typologies associated with listening preferences also have a long tradition and are an essential part of the musicology of the 20th century. Those encompass personality classifications by Besseler (1926), Müller-Freienfels (1936), Adorno (1975), Bourdieu (1982) and Schulze (1992), among others. T. W. Adorno laid a foundation with his basic listener typology, which outlines how a listener feels about music and deals with the perception and attitude towards music (Adorno, 1975). This typology results in a continuum ranging from the expert listener to the musically ignorant. Even though some of the descriptions partially align with the hereafter circumscribed types, they do not encompass the listening motivation and contextual triggers, which became increasingly important aspects for starting a listening session in a music streaming setting. As with this typology, the importance of music and music listening for the emergence of social groups with typical demographic and lifestyle characteristics was established in the 1970s (Dollase, 1986). On the one hand, there are the socio-demographic models, such as that of I. Bourdieu's class taste (1982), which are based on variables such as age or education and relies on statistical data collection for the first time. On the other hand, modern life-world concepts, such as the experience milieus, developed. Those continued the existing cultural-sociological discussion by elaborating significant, everyday aesthetic schemata and milieus that characterize the predominant experience society, as conducted by G. Schulze (1992). Another concept, referring to a listeners' environment, is offered by the Sinus Milieus (Sinus Sociovision, 2007). Those allow a further segmentation through a media user typology, a so-called Mediennutzertypologie, designed for German television. They contain holistic analyses of the living environments of the listeners, whereby value attitudes, education and income are simultaneously taken into account. However, this segmentation results in ten, partly overlapping milieus, whereas the following approach focuses on establishing a clear distinction. Nevertheless, Sinus Milieus, just as the following types, refer to parameters that take everyday aesthetic preferences and musical tastes into account and should thus allow a better adjustment of the program offerings to the expectations and habits of the audience (Oehmichen and Ridder, 2003). The therewith increasingly important music usage behavior has been thematised by J. Sloboda (2012) by integrating main activities. The following study takes this into account and aims to integrate music usage behavior by consulting listening contexts and listening motivation of individual types. This approach allows to observe behavior of listener groups with respect to their needs. This is supported by the chosen top-level model of the hereafter stated typology, whereby the removal of demographic, sociographic and genre-related restrictions allows to avoid a rapid relativization through social change. The geographical location of the examined listeners focuses on European and American listeners, which imposes the only restriction. This selection is predetermined by the predominant user base of the streaming platform Spotify during data aggregation for the analytical section.

The structure of this paper sets off with the examination of the phenomenology of music streaming in Chapter 2. This highlights the conflicts between outdated practices and recent developments in the fields of consumption, behavior, and psychology. The progressive characteristics of music streaming audiences recapitulate all unprecedented factors of those three dimensions. This summary demonstrates the need for renewed techniques in order to comprehend audiences in a manner that respects contemporary listening habits. Based on this scientific background, Chapter 3 seeks solutions to enhance understanding of the audience groups, through an individualistic process of concatenating data analytics. A methodology is provided, followed by an explanation of the analytical means used to concatenate the different datasets. A subsequent listener framework connects the insights on user behavior and preferences with the derived analytical means. This typology is designed in respect of the factors mapped out in Chapter 2, as well as the overarching need for more listener-centric insights. This is substantiated by statistical analysis, as well as a discussion and interpretation of the results. Chapter 3 concludes with examples that demonstrate the use cases of the derived audience profiles and parameter indices. Chapter 4 provides insights into how humanistic and industry disciplines can profit from a combined methodology that aims to understand audiences in the digital music experience.

In summary, this research outlines a novel method of analyzing and parsing music streaming data, with the goal of enhancing understanding of listener segments. This is supported by a renewed framework that incorporates user- and context-centric insights. The derived method of rating consumption preferences may enable the reconstruction of otherwise imperceptible, interdisciplinary correlations in a quantitative manner.

Chapter 2

Phenomenology of Music Streaming

To determine the prevalent phenomenology of music listening, the following chapter addresses the three segments of content consumption, interaction behavior, and perceptual states. The empirical research outlines recent developments and predicaments in those domains to investigate how listeners currently consume and perceive music. The interrogations present in detail why new approaches are necessary to understand audiences in a manner that respects contemporary listening habits. This establishes the groundwork for the development of the listener framework and statistical analysis in Chapter 3.

2.1 Consumption

2.1.1 Premises of Music Streaming Consumption

Today, information and communication technologies are of fundamental importance to the economy and society, as digitalization is changing the way people use media. Almost all economic sectors as well as our private lives are now heavily influenced by technologies. As a result, much of our economic growth depends on modern IT systems and digitalization, since they increase productivity and efficiency. Digital products are increasingly displacing their physical counterparts, whereby digital components are integrated into originally purely physical objects, or replacement products are designed. This process focuses on integrating all universally relevant properties of digital products by means of various functions. The six product needs, which include the demands for data centeredness, intelligence, networking, communication skills, expandability, and personalization, enable analog functions to successfully transform into digital products. Their implementation therefore aims to realize as many of the consumers' product requirements as possible. If this adaptation succeeds, an added value is created, which ultimately increases customers' benefit from the digital product (Münchner Kreis, 2016).

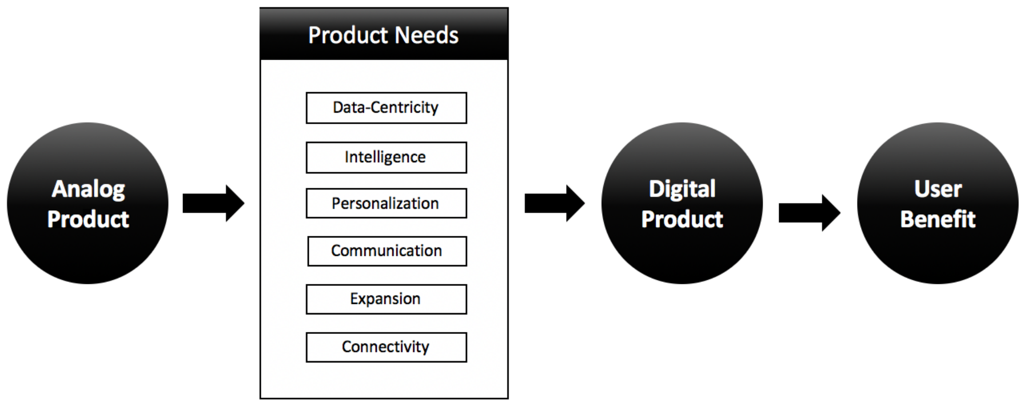

Fig. 2.1 Product and User Needs for Digital Products

Thus, it is firstly necessary to define how the standard requirements for digital media products are being met in the music streaming environment. Secondly, it is essential to know which aspects will be able to bridge the second gap, which is represented by the transition from a digital product to one that creates user benefit. Universal media and product needs have been progressively integrated into music streaming under consideration of internal as well as external functions and services. Generally, product configurations – the basic systematics as well as the dominating forms of interaction – are designed to offer the widest possible range of usage options. In particular, content delivery and its configuration have been greatly amended over the past years to meet changing consumer needs.

Data Centricity, Intelligence, Personalization

Metadata is the core that holds together the consensus of a digital file. It is data about other data, and it provides a structure and context to a given set of bits and bytes. Music metadata, more specifically, is the summary of information that pertains to a track file, such as artist name, song and album title, producer, writer, release date, genre, or track duration. Because of the more refined metadata, an extensive systematization of music catalogs is possible. This enables users to search for specific works, albums, or artists, as well as browse for moods or genres. So far, mainly data that refers to who wrote what and in what proportion, publishing data, has been curated more intensively since the switch to digital. Wrong or missing information could have negative effects on the creator's compensation and trigger non-compliance with search function regulations. The latter is also tied to a lack of personalized recommendations or playlist placements, since the metadata acts as the basis for algorithmic suggestions as well. On the basis of metadata sub-layers, usage habits, and listening histories, user-specific content can be created. In this manner, data centricity has been implemented as originally purely physical products have been equipped with digital, data-generating components.

Today, the aggregation and processing of product- and user-related metadata can facilitate new algorithms that enable a personalized and proactive approach integrating more intelligent programs into music consumption. By using metadata, the user is involved in the product design and development and thus becomes a sample consumer. This opens the innovation process and significantly shortens its cycles. In addition to the user histories, direct evaluations through feedback sessions and evaluations can be considered. An example of this is the option to favor or reject a title. Besides skipping behavior, this makes graver statements about the desired content of a public playlist. For instance, in the case of the streaming provider Spotify, those decisions surface again in the form of *Release Radar*, *Discover Weekly*, automated tracklists in radio, or algotorial playlists. To achieve this, Spotify mixes multiple recommendation strategies to develop its own unique recommendation model. These include collaborative filtering, which analyzes the behavior of a single user as well as that of others, natural language processing, for text, audio analysis, for the audio parameters of a file, and predictive analysis, to predict preferences with the help of deep learning (see

Section 2.1.2). Additional data-centric information can be found within an artist's profile in the form of concert dates, live streams of events, or most streaming countries. However, the data relevance is still limited insofar as the users cannot obtain comprehensive information about the sound recordings, because external sources must be sought for this. This is highly desirable, since references to time-sensitive topics are likely to increase the information content, which could significantly increase listeners' attention and consequently their presence on the platform. Approaches to this are the streaming modes *About the Album* and *Track by Track*, where background information about the recorded album, individual compositions, or personal insights from the artists are provided to the listener.

In conclusion, streaming services can be described as intelligent and self-determined due to their use of influenceable, intelligent technologies. These enable tasks such as the analysis and reconfiguration of user histories on an algorithmic, machine learning basis. In terms of playback control, the final selection and the playback of the works also takes place exclusively according to the will and instructions of the user, unless otherwise indicated.

Communication, Connectivity & Expansion

Updated music catalogs are ensured by a continuous renewal of its content and refreshing of alone-standing music products, such as rotating playlists. These constantly expanding music catalogs obtain their content from meta-files that are stored digitally and thus traceable as well as expandable. Today, music is from the start a cluster of data that can be stored or played on different media without being bundled. Thus, the streaming of the musical product has not dematerialized it, but made it more applicable, so that it can now be materialized by means of various products.

Users demand a product that is available anytime, anywhere, and for everyone. However, it should be noted that Spotify streaming is currently only available in 79 countries, covering barely a third of the world's total population (Spotify, 2019a). Besides the restraints given by the limited amount of countries where it is available and the requirement of having an internet connection, a wide variety of player options are available. Spotify can be used on desktop, mobile, tablet, web application, or external player devices. The use of multiple players is made possible by extensive networking. As a result, the linkage of the devices and the data transmission between them is intensified. Metadata also allows the listening experience to be controlled by multiple devices or users in various forms of human-computer interactions, including voice control. On average, all music listeners use 3.4 devices weekly to engage with music, while this figure is 3.8 devices for teens and millennials, and 4.7 devices for paying subscribers (Nielsen Holdings, 2017).

The integration of other applications into the music streaming platform has deepened the music experience just as much as it has extended the music's impact into other areas of everyday life. Such applications are Google Maps, Shazam, WhoSampled, and Nike+ Run Club, to name a few. Moreover, interactivity and social networking have increased through partnerships with social media platforms such as Facebook and Instagram (Spotify, 2019a). Thus, social media exchange, entertainment, and discourse with other users are possible. Apple contributed to this development by integrating Twitter into their music streaming platform in 2015. For some years, music recommendations communicated through social networks were considered particularly influential. Today, however, such recommendations share their position of influence with computer-generated suggestions. Nevertheless, they can be seen as an essential tool in the music consumption cycle, as 55% of all music listeners use social media to follow or stay updated about musicians they like. Furthermore, they share content and create follower bases of like-minded people with whom they connect and share experiences (Nielsen Holdings, 2017). Public or personal playlists as well as co-creations and friends' listening activities can be shared with other users. These platform activities include creation or curation and thus promote the interactive use of the offers. Respectively, in music streaming, the presentation of the information and contents must be regulated according to the listener's primary media needs. Beyond this, secondary consumer needs such as music videos, track lyrics, branding, and collaborations can further enhance the satisfaction of the product user.

Besides the six primary needs discussed so far (see Figure 2.1), A. Nylund-Hagen points out listeners' urgent requests for enhanced usability and security, among other aforementioned aspects. Firstly, a fundamental demand for data security has been raised, because across countries and age groups there is a great fear of data abuse (Nylund-Hagen, 2015). Secondly, for a large proportion of listeners, convenience is a primary value in music consumption, comparable to the former status of price and quality. According to Bay, "the parameter of convenience is determined by the attributes of the product, but just as much by the network and the context which the product is part of. [...] Convenience of a specific product is determined by its adaption to present trends, its accessibility in the market, how the product is communicated, and how easy it is to use" (Larsen, 2009, p. 30). As some listeners' need for convenience rises, their other needs decrease, such as the need for sound quality and background information about the release or artist. However, for other listeners, whose main motives for music consumption concern self-expression, mood regulation, and also identity forming and social relatedness, a need for increased convenience may not be emphasized.

In sum, the improvement of the everyday life of every single user requires more than a digital reprocessing of physical products. Ultimately, digitalization is not an end in itself but has to serve humankind (Dirks, 2016). Therefore, digitalization has put forth universal requirements for digital products and their media content. These must be further supplemented by consumer needs in the specific area of application to ultimately increase customer benefit. This ensures that the technology is by no means an end in itself, but always serves the consumer. Throughout this process, it should not be neglected that standard requirements of analog products that are not specifically attributed to the format of the end-product are still of relevance in developing a high-quality digital product.

2.1.2 Models for Product Distribution & Content

Compliance with the demands for digital products discussed above entails various options for the independent production and distribution of digital music. However, actions in these two areas are no longer exclusively being performed by record companies, as the latter have had to give up their prior position as the main driving force in the field. Despite taking initiatives to streamline and conform those processes, the distribution still differs significantly among territories. This diversity within the global music market occurs because distribution models, like content models, are shaped by new technologies of the media industry and furthermore adjust to newly emerging usage behaviors. The three major segments that constitute the media value chain are content creation, production, and distribution. Of those, we will further investigate the distribution models and the content models, since they are most exposed to technological product design. Moreover, in terms of technological developments, distribution and content strategies are closely intertwined and influence each other, whereas product creation mostly follows their directives or implements changes that can be carried out independently without affecting other links in the chain. Within the media value chain, distribution outlines the process of disseminating the content to consumers in the form of live broadcasting, media websites, and music or video streaming. Content creation, on the other hand, describes the development of original content, for instance in the form of an audio track, video, or news update.

The user's needs and the state of the technical infrastructure must be understood to assess which products fit into the framework provided by the principles of the distribution and content models and vice versa. In the following, we discuss possible start and end points for this cycle, since these have a great impact on how music is accessed.

Distribution Models

One way to create a typology of the current music streaming landscape is to differentiate between distribution models, namely, the ownership model, the access model, and the context model. These all serve autonomously but may co-exist in some market niches. Access-based services create temporary value and are possibly converted into commodities over time. However, streaming services provide listeners with access, not ownership. Thus, the question is whether access-based models can serve as long-term solutions for the music industry. Many experts claim that the economic value created by music is increasingly based on context rather than ownership. One argument supporting this is that context-based services bear a greater potential to create economic and personal value, since they generate opportunities for listeners to incorporate the act of music listening into their life, rather than only providing access to music. Thus, context models simulate contemporary versions of ownership properties (Wikström, 2012).

While configurations and deliverables still vary greatly between platforms, eventually the structure of streaming services will be the same across all providers. An ideal state of affairs presupposes an excellent music catalog and technical quality, broad territorial availability, and compatibility with all mobile devices. Access-centric services have the clear aim of making all songs ever recorded available to customers across the globe. In contrast, context-based services do not have a similarly, clearly defined target. Context-based models come in when the services provided exceed an access model by offering a greater distinction of the content. Context-based models can create unique competitive advantages by implementing innovative features, such as discovery, organization, personalization, and creative interaction, into access models. For example, in her album Biophilia, Björk uses applications to make and play with music. The provided services are less bound to general parameters of the provided content, and they expand the space through the innovative connection of the content. For instance, Spotify's emphasis in regards to service development changed around 2011 from an access-based to a context-influenced service. This change was driven by cooperations that majorly comprised social media companies, such as Facebook and Billboard. Eventually, whenever music distribution and consumption move from the physical to the digital sphere, many virtues and practices are transferred. Apple's single-track Download and eMusic's Record-of-the-month, for example, are both based on physical principles of acquisition and ownership. Yet, these two factors have decreased in importance. Experience is increasingly replacing ownership. For instance, record collections are counted as identity markers in the ownership model. In the access model, in contrast, these are replaced by the increasing importance of music listening as a social and public activity. Thus, the seemingly entrenched music identity of listeners is replaced by a continuous flow of information describing their real-time musical experiences.

Another way to differentiate the models is through their similarities with goods and services. For example, ownership models are comparable to goods in product-related industries, such as shoes or books. In contrast, the access model is often marketed in a service-centered way, where services are being packaged and sold, such as in hotels or banks. This shift has been majorly induced by novel consumer behavior. Music is gradually transforming into a service rather than a product, where service providers must always fulfil basic product needs and offer complementing services. Context models have to adhere to the digital product needs and furthermore cater to user needs as individually as possible within the given sector. All six universal product needs exist and are relevant in digital music consumption, but their implementations are developed to varying degrees and diverge particularly in the areas of computer-human interaction. Innovators who manage to derive market insights from pioneering technologies will be able to provide product platforms that correlate with actual consumer needs (Wikström, 2012). New listening experiences that go beyond access-based services include personalized playlists and recommendations, among other elements. This all requires user-specific data on listening habits to surpass the standard configurations of an access model. However, this data is being used not only to nurture a new distribution model, but also to utilize content for upcoming elements by leveraging those insights with third-party providers. For instance, Spotify's data on streaming consumption and interaction behavior is being employed by labels, artists, and promoters. Within this context, the information has a bearing on conceptualizations that go far beyond distribution strategies. For instance, album releases, artist collaborations, and concert tours are being based on these insights. As platforms grow larger, it can be assumed that the impact of their accumulating data will gain in value and increasingly facilitate the creation of new features and content elements (Münchner Kreis, 2013).

Although more context models are most desired, certain side-effects need to be kept in mind. As publications of the Foundation for Future Studies point out, the population has a noticeable tendency to feel overburdened by media in everyday life due to the constant presence and flood of information. A central problem is the differentiation of important from unimportant information. Accordingly, there is an ambivalence, since permanent availability is perceived on the one hand as progress, and on the other hand as a limitation and burden. Hence, aspects such as information overload must be contained by systematizing the media contents, by filtering data sets as far as possible in advance (Stiftung für Zukunftsfragen, 2018).

Content Models

The ultimate aim is to attain a mix of content formats that cater to the access model and that likewise, with the future in mind, could also cater to the context model. All services and functionalities on music streaming platforms are facilitated by access content. Hence, *context content* builds upon this structure to increasingly satisfy contextual needs, which are consolidated in the base framework provided by the elements comprising the *access content*. Thus, both content options are necessary to meet the customer's expectations. However, access elements are the ones referred to as ubiquitous and standard elements, whereas context elements serve to satisfy the increasing demands.

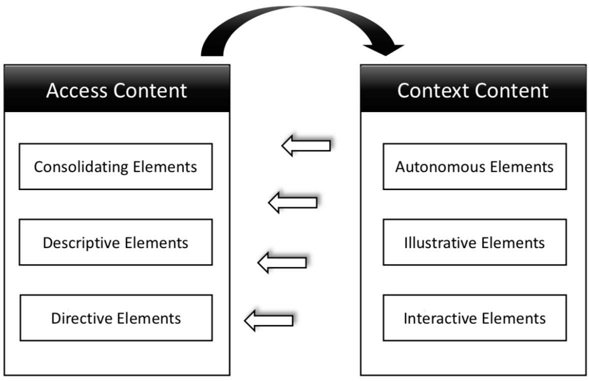

Fig. 2.2 Content Models: Access & Context Content

Neither category should be seen as static. From a broader perspective, the primary objective of streaming providers is to shift their focus from an access to a context model by upgrading the corresponding content elements. Simultaneously, from a narrower view, the increasing demand for context content induces a backward-oriented migration, whereby former elements of the context content are being established as access content. This occurs once a content element is fully integrated not only into the platform but also into the user's interaction with the content and recurring usage processes. Thereby, the migration and correlating change in naming conventions is directly connected with the commonality and ubiquity of certain elements. In this manner, for instance, one extending element that was once a context element is now an access item.

The dynamic structure of the value chain in the music industry is causing the major labels, which have dominated the production of media over the past decades, to diminish in power. This process quickened during the rise of digital distribution, when decreasing production costs for marketing and distribution turned the product features and deliverables upside down. The accompanying distribution with intangible mediums enables the creation of new ways to connect artists with end-consumers on various levels by implementing digital technologies. This disruptive technical knowledge incorporated by digital music is mostly emerging outside of the major players' existing value network. However, it is exactly these alien elements that have stimulated the creation of alternative formats and elements for music consumption (Larsen, 2009). The contributing factors for those newly emerging elements reach beyond music-centric attributions. F. Holt argues that musical practices must be observed in a broader context, including social and technological changes, to assess their value (Holt, 2010). This demonstrates that the music industry is successively moving away from an access-based model towards a context-based model (Wikström, 2012).

Access Content Type	Components
Consolidating elements:	artist information, cover artwork
Descriptive elements:	sound metrics, album title, track title, artist, contributing parties, release date, label, track identifier, genre, mood
Directive elements:	search, filter, playback options, library saves, playlist creation

Table 2.1 Access Content: Elements & Components

The access content is comprised of three element categories with its unique components. The directive elements comprise navigation features such as search, filter, and various playback options. The latter include play, pause, skip, shuffle, crossfade, library saves, as well as downloads and the option to create playlists. These enable the user to exert direct control over the search and play process within the platform. Firstly, however, finding and differentiating the music files' descriptive elements play a crucial role. Metadata is one of the key drivers of the operations of major music streaming services, and it represents an integral part of today's music experience. In the case of Apple Music, Google, and Spotify, the content for this data-driven substructure is provided by Gracenote, which in 2019 is the industry's most comprehensive source of descriptive music data. Gracenote's backlog of descriptive elements encompasses roughly every song ever recorded and includes, besides the descriptors for the music dataset, further aspects such as the genre of the song, the era in which the song was recorded, the origin or region most associated with the artist, the language of the artist, the artist type (mixed, female, male), the mood (e.g. rowdy, somber), the tempo (e.g. fast, BPM), and the style (e.g. Industrial, Jump Blues). In total, Gracenote's descriptor system contains 2,451 genres, 438 style descriptors, and 480 languages, and these are utilized to make deep connections between artists and tracks, creating radio stations and playlists that share common musical characteristics (Gracenote, 2019). Without the provision of this global music data to streaming platforms, a listener would not be able to listen to, find, or even see a track on a streaming platform due to the missing access content.

Directive, descriptive (audio file level), and consolidating elements represent the core elements for audio playback. The console of the Spotify Web API lets users explore those foundational data points through an easy-to-use interface. Users have the option to enter anything they wish to find, ranging from songs, albums, and artists, to playlists, podcasts, videos, genres, moods, year, label, and ISRCs and UPCs. Those are considered to be standard elements and are omnipresent prerequisites on streaming platforms. With the ingestion of an audio track, the platform extracts all audio related sound metrics. Spotify allows users to have a look at these as well, including acousticness, danceability, energy, instrumentalness, liveness, speechiness, and valence, among others. These sound parameters in combination with the metadata elements allow the systematic organization of the music files. They also facilitate the categorization into thematic or genre batches, which streamlines the curation of playlists. All in all, the Spotify API allows the yielding of detailed information about a track's audio features (Spotify, 2019b). Additional consolidating elements contributing to the basis of the access content include background information on the majority of artists as well as cover artworks, which constitute the single visual indicator among the access content elements.

Context Content Type	Components
Autonomous elements:	personalized playlists, recommendations, playlist queue, history
Illustrative elements:	videos, virtual reality, lyrics, about-the-track insights, related artists
Interactive elements:	live shows, collaborations, friends' activity, content sharing

Table 2.2 Context Content: Elements & Components

Every content element that lies outside the scope of primary music listening and navigation will first and foremost be described as a context element. Context content also consists of three entities, namely interactive, illustrative, and autonomous elements. One of the main benefits of a context-oriented approach is that it enables consumers to experience music rather than just listen to it. This is mainly driven by the development and integration of new tools that add value in the experiential field of context uniquely defined by each user. This requires a significantly higher volume of data than the access elements. However, as novel features require growing volumes of data, more infrastructure, and intelligent networks, elements are firstly referred to as context content and possibly migrate at a later point into the access section. This also implies that context-based models only emerged once digital capacities and technologies were integrated into music streaming platforms. Among these elaborate features are personalized playlists such as *Discover Weekly* and *Release Radar*, *Spotify Radio*, *Daily Mix*, and stations, which are all based on a recommender system. Further extending elements contributing to the context models are key lyrics, related artists, and discovery features for playlist queue and history, friends' activity, and content sharing.

While personalized services are becoming increasingly important, consumers' influence on content and its mediation has grown in proportion. Accordingly, consumers, also called prosumers in extreme cases, strongly influence their music experience. This is due to the use of their own usage history, which is collected during interactions with content and later repurposed for content refinements. The influence on the listening experience can be enhanced by a high standby, activity, and interaction potential, as higher degrees of interaction allow for vaster collections of usage data. Moreover, autonomous elements assist in overcoming biases by weighing active ratings and statements about wishes against a listener's factual user behavior (Brüggemann, 2017). Consequently, this intelligence of the underlying system must be considered in the process of product creation, to prevent restrictions and to be able to offer an open system where users can switch situationally between content and curation options. This is supported by individual elements of the product and media design, especially the various playlist configurations. The relevance of this playback option can be explained by the combination of lean-back- and lean-forward-oriented functions (see Chapter 2.2.1), which enable the music streaming experience to be individualized and ultimately intensified for each user type. As a result, the digitized music catalogs, which were once created for their own ends, have evolved into a value-added product through intelligent systematization.

According to Robert Prey, "In an age of personalized media, the word ‘masses’ seems like an anachronism. Nevertheless, [...] there are in fact no individuals, but only ways of seeing people as individuals" (Prey, 2017, p.1086). Music taste provides a large platform to analyze someone as an individual, since it reflects memories, aspirations, everyday life, and social circles. In the age of personalization, the distinct taste and predilections of an individual listener are observed and ultimately make up measurable types that are aggregated into profiles. Autonomous elements are reliant on exactly those user-centric insights to provide music recommendations or music discovery. These two umbrella terms usually refer to one of the following things: artist or song similarity, personalized recommendation, or playlist generation. Recommendation strategies are mainly driven by four music knowledge approaches with which Spotify creates its unique recommendation model. Those include collaborative filtering, natural language processing, audio analysis, and predictive analysis, as already circumstantiated. Those features allow to predict preferences with the help of deep learning.

Firstly, collaborative filtering identifies which users’ musical tastes are most similar to those of an individual by comparing that individual’s vector with all of the other users’ vectors. This ultimately indicates which users are the closest matches. The same occurs with the Y-vector, which filters for songs. *Discover Weekly* is built upon the process of collaborative filtering: the track lists of these playlists are constructed based on the user’s own listening history, paired with the listening history of people who display a similar taste. Secondly, the results are further substantiated by natural language processing. This regards content-based recommendations, whereby the source data for these models contains words in the track’s metadata, lyrics, news articles, and social media platforms, among others. This ability permits to harness human speech and close the semantic gap between music audio and various surrounding aspects. Thirdly, audio analysis examines the song’s key characteristics, which helps to understand fundamental similarities between songs. For instance, *Spotify Radio* is based on audio analysis, which is conducted by scanning a seed track, artist, playlist, and similar music for their sound attributes and genre. Moreover, in time, a seed track facilitates discovery based on one track a user enjoys. The main aim of this recommendation strategy is to give listeners the ease of receiving a pre-curated playlist of tracks that they might want to try based on a given seed. Fourthly, in predictive analysis, convolutional neural networks are used to recommend music. The inputs for this are time-frequency representations of audio signal frames, which are concatenated to form a spectrogram. Predictive analysis facilitates the detection of a listener's preferences with deep learning. Hybrid playlists and recommendations are part of this category. Overall, an increasing number of music streaming features rely on predictive analysis. This enables, for example, the compilation of *Daily Mix* playlists, which are based on styles someone listens to and which mix tracks for new discovery with already known tracks in a personalized playlist. The aim of this product is to divide a listener's tracks into genres while adding some new tracks that fit into the mix for discovery. Predictive analysis also underlies *Release Radar*, whereby a mix of new releases from the artists one listens to most, top artists and followed artists, are mixed with unknown tracks. Those tracks are majorly determined based on artists listened to as well as liked and disliked tracks (Spotify, 2018). Based on the four recommendation algorithms discussed above, the instrumentation of user-specific metadata will continue to grow in the future due to the planned integration of deep learning, artificial intelligence, and virtual reality. The extent to which these innovations will expand the utility for the music consumer, or whether the products will instead be subject to a technological end in itself, is continuously being evaluated within prototype cycles (Brüggemann, 2017).

Lastly, the category of context content encompasses illustrative elements such as artwork. To users, those indicate the possibility for discovery, inspiration, and the choice of a product based on not only content criteria but also visual components. "As multi-sensory creatures, the act of choosing what music to listen to has—for a very long time—been intrinsically tied to an aesthetic experience" (Stocks, 2017, p.1). For this reason, artwork belongs to access content. However, this element is entirely missing from AI-powered music experiences and voice-activated smart speakers; for example Amazon Music listeners can build playlists via Alexa exclusively using voice commands. Accordingly, other illustrative elements are filling this gap. Decisive factors for the future embedment of virtual reality and on-demand video are their social and immersive aspects, which are essential to current users. This way, listeners with similar preferences can watch content together in surroundings such as a virtual reality living room. The audience thereby has the freedom to look anywhere within a music video or experience concerts together with other fans, as if they were physically present (Ericsson Consumerlab, 2017). Likewise, to showcase how audio-visual supplements can capture an audience's attention, D. Bakula, SVP Client Development at Nielsen Entertainment, draws attention to highly popular music videos such as PSY's Gangnam Style on YouTube, from which people replicate the dance moves. Spotify integrates such features by offering visual elements to complement podcasts (Johnson, 2018). This format is called *Spotlight* and focuses on music, pop culture, and politics (Münchner Kreis, 2013).

Beyond that, prototypes show other technological advancements that could potentially also be integrated into next-generation music streaming platforms. Spatial computing has the power to blend technology into the real world. This way, virtual reality, augmented reality, and mixed reality have the capacity to distort the frontiers between live and mediated performances. Such features tie into the interactive elements that have been enabled by the pairing of the diversity, ubiquity, and availability of content with a powerful user, which has impacted the current media and information landscape (Münchner Kreis, 2016). This original and personalized content ties back to users' need to express themselves by making music a part of their social identity. Therefore the cooperation with social media platforms became of increasing relevance in order to support a deeper engagement between artists and fans. The synergy between music and community building as it opens up creativity, connections and innovation through music and video. One of the central social features on Spotify is *Friend Activity* which shows the listening activity of friends as well as public profiles, such as brands and celebrities, that you follow. Moreover, for the first time, Spotify's new feature called *Social Listening* allows multiple people to add songs to a queue to which they can all listen. Innovations are capable of creating indistinguishable boundaries between live and mediated sessions, as influenced by sound metrics, or between the production and consumption of music, as in the example of co-creation. According to S. Vargo and R. Lusch, this service-dominant logic of marketing co-creation entails that engaged consumers are a part of the creation and attribution of meaning to products, services, and experiences (Vargo and Lusch, 2008). According to T. Ramsey White et al., this interaction means that "co-creation is an integral part of the artistic experience, where audiences engage in cognitive, emotional, and imaginal practices to make sense of the performance" (Charron, 2017, p.2). This way, consumers are no longer just passive recipients of the content, but they can instead take on the role of co-creators (see Section 2.2.1).

In sum, while access and context distributions can be applied separately, their context entities are amalgamated. A platform aiming at context content would not be able to organize or display its content, since context content relies on access content at all times. Therefore, all access content needs to exist first before any context content can be built on its foundation.

2.2 Behavior

2.2.1 Premises of Listening Behavior

The most prevalent insights on a listener's action can be detected by observing the level of focus, actionability, and receptivity (Ross, 2010). On the one hand, the intentions and desires of an exploratory and interactive *lean-forward user* reinforce those three aspects. On the other hand, *lean-back users* ground their cognitive and behavioral patterns on a model that demands a minimum of mental and physical effort. The key elements by which these two types of users can be differentiated are their physical interaction and their mental awareness.

Since physical record sales are declining, the focus of the industry is increasingly shifting from the sale of units to the playback of recordings. It must be observed that time and attention are scarce resources in the digital space. However, exactly those are required to anticipate longer interactions with the presented content, which are essentially the determinants of financial return. The transformation of the distribution and content supply on music streaming platforms has yielded two different approaches to increase listener engagement while catering to different listening needs and behaviors. The first approach uses a variety of navigation features that facilitate access to a vast amount of content. The second takes the form of new applications and interaction features that enable and encourage the user to shape the listening experience hands-on (Münchner Kreis, 2013). Both of these redesigned approaches involve upgrades targeted at both listening types, namely lean-forward and lean-back users, to cater to the needs of listeners at both ends of the spectrum.

Lean-forward users are characterized by higher than average degrees of focus, actionability, and receptivity on a physical and mental level. Thus, such listeners represent a desirable client profile for music providers, since they are more likely to consciously and actively engage with the provided content. The addition of interaction and discovery modes is particularly linked to lean-forward users and can be depicted by a change in terminology in common parlance: namely, in the streaming context, music consumers today are not only identified as listeners, but also as users. Accordingly, the term that describes the music consumer has changed from having a passive connotation to an active one. The conscious utilization of the descriptors listener and user indicate the degree of activity displayed by consumers. This is reinforced by an increasing number of usage options, such as access services or algorithmic recommender systems, that allow for the user's individualized handling of the systems.

Furthermore, the necessary participation can be scaled up to the form of a prosumer. A prosumer is a user who plays the roles of both the producer and the consumer. This is the case, for example, when a blogger reads other bloggers' posts and at the same time contributes content to the platform by commenting on those posts. The same applies to users of streaming portals who consume music and share it on social networks or create their own playlists. According to S. Miles, "In effect, the individual consumer becomes the conduit for his or her own consumer-driven definition." He argues that this meets the needs of a prototypical modern consumer, who is an “inevitably disappointed authenticity-seeker” (Miles, 2018, p.23). Thus, the evolution of consumers to directors of their experiences becomes evident, as listeners are in motion, rather than the music. This leads to multifarious variants of the music listening experience instead of one single experience that has been determined by the industry (Miles, 2018).

More self-curated listening experiences were reported in 2017 than ever before. This change has been promoted by the diversity of options in playlist curation, choice of device, and connection of digital profiles. Apart from standard playing features, self-curation of music content and playlist creation are heavily used interactive features that enable users to have influence on their choice of playlists. A study by Nielsen on listeners in the United States highlighted that 58% of listeners create their own playlist, and 32% share their playlists. Furthermore, 38% of all streaming listeners agree that playlists are an important part of their streaming and experience, and 48% of those prefer to curate their own playlists over listening to other playlists. It seems that the very lack of materiality, ownership, and emotional resonance on streaming platforms motivates users to curate and arrange audio files themselves. This leads to a listening experience that strives to create something tangible to enhance the perception of the musical medium in streaming (Nielsen Holdings, 2017). All those playback options are intended to garner consumers’ attention and interest, and they simultaneously increase the number of titles per listening session, which overall results in longer interaction periods on the platform.

In contrast to lean-forward users, lean-back users are characterized by lower than average degrees of focus, actionability, and receptivity on a physical and mental level. Thus, these listeners have different basic demands than their counterparts. Although lean-forward users are favored for their higher awareness factor, the lean-back elements represent an important resource for music providers, especially in regards to contextual properties, which were not been fully tapped until recent years. While providers supply an increasing number of product ranges, the exploration and the needed familiarity with the titles depends more than ever on listeners’ personal initiative and knowledge. Automated playback options and computer-generated recommendations offer content to this user segment in the most convenient way possible (Kachkach, 2016). Since those applications set up an environment where musical engagement is no longer directly linked to actions taken by a listener, the process of listening becomes increasingly passive, requiring only a few motor and mental actions. This demands less readiness to act and physical actions from the listener, and thus allows for a relaxed listening process. However, while the listening processes tend towards more passive engagement, more opportunities exist for active agency than ever before. Therefore, intensified experiences can occur despite a minimalized and streamlined frontend environment. This is because the chances that someone will interact with some content in the music catalog are higher when multiple playback options are provided for scenarios in which they cannot control the session or are overwhelmed by making a selection. Those options are especially utilized by lean-back users, who display a very low level of motor and mental activity when streaming music.

This type of listening behavior is also known to accompany various everyday situations. For instance, music may be playing in the background during activities such as sports, cooking, or tidying up. This way of consuming music is often characterized as superficial. This originates in the perception of encountering a commonplace context, when music listening is extensively practiced as an accompanying activity. Yet, listeners may evaluate the exact same listening experience as intensified. This may occur because music played in the background bypasses boredom, creates a new atmosphere, or makes tedious work easier. As a result, the semantics change due to the temporary motivations and activities. The resulting level of attention can usually be deduced from the current inner and outer situation of hearing and determines the hearing behavior. The listening context is composed of factors such as the time of day and the social environment, as well as the individual's mood and activity. To address listeners in a more targeted manner and thus gain more of their attention, daytime-specific adaptations to playlist rotations, among other things, are made, so that they form a dynamic structure. Thus, it is possible to take daytime activities into consideration which shape everyday situations with their specific dynamics (Nylund-Hagen, 2016).

The increasing demand for music in everyday scenarios has led to demand for new navigation options, as becomes apparent when observing the most frequent search options on Spotify and Google Music. Today, it is uncommon for users to employ bibliographic terminology when searching for music on access-based music streaming platforms. Instead, they tend to query descriptive categories, such as emotions or context. Thus, a large variety of search options have been developed to adapt to those new query practices. To illustrate this shift from content-related to context-related music content not only from a search but also from an inventory perspective, Chartmetric analyzed and organized all Spotify playlists based on context (CX), content (CN), and hybrid (HB) purposes. Content-based playlists are grouped based on the track genres, language, or geographical boundaries, such as K-Pop Acoustics or Today's Top Hits. Context-based playlists can be activity- or time-related – for instance, Running or Deep Focus. According to them, CN playlists continue to account for the majority (57%) of the genres & mood playlists, which are followed by 211 Mio.

listeners. In this manner, genre and music eras, still seem to be the favored navigation help to find songs. However, this might be a remnant of the past due to the duration of the dominance of this playlist type and the formation of the others. This notion is further strengthened by looking at the median number of followers as well as the follower gain, it is clear that hybrid playlists lead the game, followed by context playlists (Joven, 2018)[1]. This change in search terminology and content naming has been enforced since retrieval systems neglecting musical, cultural, or personal aspects increasingly risked becoming obsolete for the contemporary ways of dealing and interacting with music.

Key elements of this transformation were compiled by the area of research known as music information retrieval (MIR). MIR evolved in response to challenges and specific needs in this domain in the 1990s, bound by the International Symposium for Music Information Retrieval. The main mission of MIR is to extract descriptors from audio signals or contextual sources that are meaningful to the music listening process, to improve the retrieval, browsing, and recommendation of music content. Those three elements can be differentiated by the user's intention, as per retrieval of specific music content, browsing for unspecified content, or allowing the system to recommend potentially relevant items based on actions and preferences. According to P. Knees and M. Schedl, in general it is the aim of feature extraction to transform raw data that represents the music item. The result should be a more descriptive representation, describing musical aspects as perceived by humans. This could for instance be a paraphrasing of heard instrumentation or harmony that is easy for a listener to recollect (Knees and Schedl, 2016). Such intelligent retrieval systems enable a lean-back user to receive customized recommendations as well as better search tags that are more appropriate for those thinking in contextual attributes. In addition, puristic, easy-to-understand, and uniform interfaces have been designed to meet the product requirements of these consumers.

However, before a listener considers how to access the desired content, the stimulus threshold has to be met. This describes the level of activation necessary to induce an action and can best be exemplified by the playback mode of vinyl records. Due to the related physical action, this medium always requires the consideration of whether the needle should be taken off the vinyl or not. If the listener only wants to skip three minutes of a recording, this desire often does not seem proportionate to the required effort. Hence, ordinarily, the listener completes the entire album or most of the tracks on a record. Accordingly, the titles are always presented and played in the album context. The audio recording is designed as an artistic unit where the title sequence is permanently fixed. This playback and listening patterns peaked in the times of CDs and vinyl records but is no longer existent in the streaming

[1]Analysis timeframe: March 2017 - March 2018

age. The play option that is closest to the former mode is the option to search via the album section, whereby a title from an album context is selected first. Today, there are numerous ways to intervene in the predetermined play order. The active usage of these interaction options characterizes lean-forward users, because they access those options recognizably more often than the lean-back users do. The selection is made between direct search, shuffle mode, and skipping. Skipping is the most commonly used option and describes the switch to the next title (Lamere, 2014). Whether by shaking, speech, or pressing a button, the next track is not far away. Thus, a user can cause self-dynamics of playlists by acoustic, tactile, or motor means, even if they are created as static playlists. As a result, users are never tied to the given order: they always have an opportunity to intervene. If one compares the number of skips under the individual search options, one notices that a high skip frequency occurs especially with the *Discovery Features*, whereas with the *Familiarity Features*, like artists, known pieces are targeted, and a lower skip frequency is therefore recorded (Kachkach, 2016).

Regarding the timing of a skip action, the externally created features, such as radio or public playlists, show the shortest stamina before skipping a track. On the other hand, in the targeted search, in which an already more strongly filtered selection is available, longer consideration periods prevail before a skip occurs. According to P. Lamere, an expert in music technology, active lean-forward teenagers use the skip option most frequently, whereas older generations use it less than the average. In total, more than 82% of all streams are either listened to completely or skipped within the first 5% of the track. This suggests a very deliberate choice, as well as the prior relevance of the first seconds of a recording (Kachkach, 2016). The highest probability of jumping to the next track before the end of the track is within the first 10 seconds (Lamere, 2014). Such interventions are always dependent on user behavior and activity and can therefore diverge greatly between lean-forward and lean-back users. This can be explained by the fact that the psychological barrier for active interaction with the system during the playback process is lower for a lean-forward user. In contrast, in most cases, the potential of the stimulus is too low to trigger an action for a lean-back user. Accordingly, this group also shows numerous track-down phases, with multiple tracks being played in full length and in their predetermined order.

This topic is further broadened when one considers that lean-back users often casually and repetitively listen to music for several hours. As a result, the acoustic event becomes a kind of wallpaper music. A. Kassabian elaborates on this with the idea of omnipresent hearing, describing this state as a “notion of ubiquitous listening” (Kassabian, 2013, p.18). This term refers to the act of hearing while simultaneously performing another action, whereby it is not clearly defined which action takes up the primary or secondary position. Such incidental listening occurs when the listener’s attention is captured by the acoustic structures but only for a brief moment. This type of streaming is accompanied by a usage dominated by successive track-downs. A track-down refers to playing multiple tracks in their full length and in their predetermined order, and it occurs, for example, if one runs a playlist without intervention (Kassabian, 2013). In some special cases, such as in the classical genre, low track-down rates are noticeable. This cannot be justified solely by an attention deficit, as among classical consumers, an above-average number of people deliberately take time to listen to one or two albums completely, and these albums encompass tracks that have above-average durations. In contrast, click-rows are for the most part induced by lean-forward users, who deliberately select singular titles or instead skip to any title in the queue. Thus, a lower interaction potential is required for track-downs than for click-rows, which illustrates the difference in activation levels of both user types.

Based on this, the question arises of whether casual hearing can be described as passive listening. According to the thesis of style historian H. Besseler, the passive listener is someone who is facing an event and expects to be carried away by it without being urged to participate in its realization. In this way, the listener still internalizes the sounds, which enable him or her to have feelings. From a philosophical perspective, H. Plessner describes passive feelings as the perception of the beauty of music, in addition to its determining power: Detached from its acoustomotor and sensorimotor behavioral context, listening becomes an independent state of consciousness. In F. Nietzsche's view, listening is a floating rather than dancing sensation (Plessner, 1980). With this, the habitus of hearing aims to transform the acoustic and sensomotoric functions into a new form of sensation. Thus, one can describe as passive a state in which the link between the hearing process and the subsequent reaction are disregarded. In this state, acousto-motor triggers are suppressed, whereas they are expressed by active users in forms such as dance. The normal cycle of the sound reception is disturbed by the decoupling of the recipient from the sound-producing units of the ear.

The conclusion is that the term of passivity in the listening context should not be related to a holistic lack of physical and mental activity (Krüger, 1999). Even in sleep, such a condition can never occur because of respiration and brain activity. Even after the removal of any mental control, the body will undergo physical experiences (Herrmann-Sinai, 2009). In this respect, the terminus of passivity in the listening context can only be validated if one changes the approach and considers it as an umbrella term for all activities that are not intentionally controlled, without the connection of the hearing process to the reaction. D. Vaitl defines it as a specific psychological process that moves on a continuum of activation-deactivation to the pole of a fictitious basal value and is characterized by feelings of well-being, calmness, and relaxation (Vaitl and Petermann, 2000). Although interactions might be kept to a minimum by one listener type, listening to music always requires a certain amount of activity. For example, despite personalized pre-selection, the user's first selection of the music cannot be completely removed at this point. While both listening types are clearly differentiable in their extremes, as in the above examples, transitory zones must be kept in mind. The listening process is subject to many outer as well as inner influences, so it is necessary for such a system to be open and flexible for user requirements.

2.2.2 Synergy of Behavioral Patterns

While lean-back and lean-forward users prefer different play options and use them with different degrees of activity and attention, they should not be regarded as two independent user types that are present in a fixed form. Instead, fluid transitions exist between the two, resulting from inner and outer influences discontinuous from the streaming act as well as aspects motivated by the choice of playback feature and music content. Therefore, the user-frontend, with its navigation tools and the interrelation between different playback modes, must be versatile to create a tool capable of accommodating situational changes.

Properties in which both listener types operate very similarly can be observed in the sphere of mental and physical activation potential. According to H. Motte-Haber, for both types of users, maximum satisfaction is felt when there is moderate excitement and complexity. This means that with a rising perceptual threshold, the level of activation increases, which leads to an immediate decrease in pleasure (de la Motte-Haber et al., 1996). However, this thesis by H. Motte-Haber can be countered by the fact that both listener types have different scopes of perception, for example in regards to boredom or stress. In some cases, extreme sentiments can be perceived as positive. Thus, even with this definition, no clear limits are imposed on the reception and perception of the music, as these are defined by the inner aspects of the consumer as well as the external aspects of the surroundings. Accordingly, the lean-back and lean-forward states should not be regarded as static.

In this sense, H. Kornhuber and L. Deecke showed that a basic potential of readiness is always available. They conducted a reverse analysis of the neuro-electric processes that accompany repeated, arbitrary movements. Their analysis explains how the readiness potential, also called surface-negative brain potential, grows with the intentional participation of the subject, while at the same time the level of attention increases exponentially. In the case of indifference, however, the level of attention is lower. Since the standby potential does not diverge in the same way, these authors' reverse analysis examines how this differs between the two forms of hearing. S. Kirschner and M. Tomasello's investigations are consistent with the main results of the reverse analysis. Both endorse that before lean-back or lean-forward-oriented motions set off, there is at all times first a surface-negative readiness potential, which later changes largely to the same extent (Kornhuber and Deecke, 2010). Thus, an activating and motivating readiness potential is a given at all times. These two factors are both essential to users who want to use on-demand offers. Consequently, a basic level of activity can be demanded from both lean-back and lean-forward users. Both types are defined as being neuronally prepared, so that from the bottom up neither of them is fixed, but can only be defined based on the actions they take.

Once they starts to play music, lean-back and lean-forward users are more clearly distinguishable. During the playback process, features vary in the required degree of activity needed to reach the desired selection or listening channel. Particularly in the period between accessing the platform and playing a title, more cognitive work may be required depending on the chosen navigation and playback element. Lean-back users allow the music to play back without interfering, while lean-forward users have an interest in actively shaping the listening process. Hence, this type of user will take advantage of extending content elements among the access content (see Section 2.1.2). In contrast, the autonomous elements among the access content, including playlist configurations and recommendations, support the listening process to such a high degree that lean-back users can obtain a music product that displays high diversity even without any intervention on their behalf. During the playback, the play modes with their Familiarity and Discovery Features represent indicators for differentiation. Lean-back users tend to use the Familiarity mode, while lean-forward users prefer the Discovery mode. However, those features cannot be unanimously assigned to a type, as they only correlate with the generally preferred low or high activity levels of these users. There are no absolute requirements, because a lean-back user can in some cases also select a discovery feature such as the radio option and play it without intervening. For this, no adaptation to the lean-forward character is necessary, although it is questionable whether the stimulus impulses remain so low throughout a rotation applied to this type of exploration that a user characterized as lean-back will not nevertheless record a higher activity rate than usual in this situation. The playback modes provide a general orientation of the playback options with which these two types of users feel most comfortable. Users will also notice such an assignment if they consciously consider the choice of a playback feature and content element. This indicates that over time, lean-back and lean-forward users are subject to a fluctuating character. This reflects one of the freedoms that users experience in the streaming environment. Unlike physical products, users have the ability to determine in minute detail what content they want to hear and in what compilation. This can be shaped up to the last moment when the recording is played. This transitory character for instance also allows users to switch without effort between perceiving music as an auditory element in the attentive foreground and in the inattentive background state.

Especially in the presence of not only those two behavioral patterns but also a transitory state, the division of preferences in listening behavior seems at times trivial. However, it is essential to accentuate cases in which a creator, owner, or distributor of music content would rely on an isolation of a specified listener type. This occurs when one wants to understand in what scenarios music is being perceived as an element in the foreground or background. It can also facilitate the strategy when identifying listeners who engage actively and knowingly with the music and the content formats that are most suitable for them. Furthermore, playlists can help in splitting the audience as per preferred behavior in certain contexts. This dissection enables a contributor to understand the respective groups of listeners on a basic level and furthermore to detect relations between music content and listening contexts.

2.3 Psychology

2.3.1 Premises of Emotional Perception

The power to represent or express meaning is one of music's most pervasive features. In particular "the ability of music to represent or express emotions" (Cespedes-Guevara and Eerola, 2018, p.2) is one of the main cause for it's omnipresence as well as it's motivational power. The perceived emotions are based on the features that are embedded in the music, and those need to be decoded to understand the feelings that lead the listener to a specific emotion. So far, research studies have not dealt with music's variety of emotional and non-emotional meanings in everyday contexts. Recent research on the categorization of emotions had led to some controversy regarding how best to describe emotions in relation to music. To date, many music-emotion related research projects have been conducted, though with a multitude of approaches and settings all covering singular dimensions of the research question. This highlights the need for a revised approach (Nawaz et al., 2018).

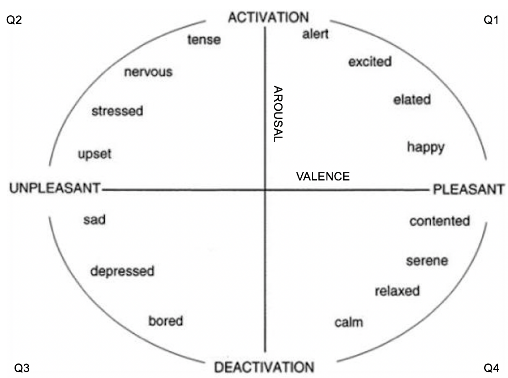

Fig. 2.3 Circumplex Model by J. Russell (Posner et al., 2005)

Currently two models are predominant in the space of music perception. The first model emphasizes "music's ability to express a limited set of so-called basic emotions" (Cespedes-Guevara and Eerola, 2018, p.2), which usually include the categories of happiness, sadness, fear, and anger. At times, additional divisions such as tenderness or love are subjoined. The second model is J. Russell's circumplex model, which measures musical expressivity by means of arousal (active/passive) and valence/pleasure (negative/positive) as basic dimensions of affect. However, these two dimensions are not capable of exhausting all affective specifications that musical material can provide (Cespedes-Guevara and Eerola, 2018). Energy and tension are two replacement factors to which listeners are also sensitive. For example, G. Ilie and W. Thompson[2] state: “The most important difference between a dimensional and a categorical approach is that the former assumes that emotions vary in a continuous manner in ’emotion space’, whereas the latter assumes that there is discontinuity (discreteness) in ’emotion space’” (Juslin, 2013, p.3). The following paragraphs ought to challenge that musical expressivity is presumably organized around a set of discrete, basic emotions; a constructionist account is proposed instead. The main concerns with current perceptions are based on conflicting points in three areas of music perception.

The first is the issue of moods versus emotions. As stated above, among the basic emotions, five categories are usually included: happiness, anger, fear, sadness, and love-tenderness. Other categories, such as disgust, contempt, guilt, shame, and lust, are neglected. This allows to conclude that the emotions most frequently brought up in music research describe affective states which do not require an intentional object. In contrary, categories like "disgust, guilt, shame, and lust are always intentional states" (Cespedes-Guevara and Eerola, 2018, p.5), since they are directed to an object. For example, if one feels guilt, one feels guilty about something particular. This brings up the question whether both types of reception, object-bound and -unbound states can be expressed by music. This can be assessed by looking at the phylogenetically inherited character of emotions and moods. For emotions that is "quick, object-directed, motivationally driving reactions, moods [are in contrast] slow, diffuse, cognitive-biasing states" (Cespedes-Guevara and Eerola, 2018, p.6). There seems to be consensus that the dimensional approach focuses on subjective experience such as moods, and especially feelings, while it performs insufficiently in accounting for emotional expressions. However, it is important to keep in mind that dimensional models are derived from “abstract dimensions that resulted from multivariate statistical techniques applied to similarity ratings of facial expressions and emotion labels” (Juslin, 2013, p.4). In that respect, another ongoing discussion should be considered: whether there can be one cognitive mechanism that processes emotional and non-emotional stimuli, or whether emotional stimuli require a separate mechanism that focuses entirely on those stimuli. On the one hand, if researchers evaluated emotion and cognition as autonomous, disconnected psychological processes, only few insights on emotional perception could be gained when examining cognitive mechanisms of perception. On the other hand, if emotional and cognitive processes are merely categorized as different classes of stimuli that are being processed by the same mechanism, cognitive mechanisms can deliver additional insights on the perception of emotional stimuli. This approach is substantiated by the fact that even though emotional stimuli are slightly more complex and emotional quality takes advantage of higher priority processing within the cognitive-affective system, it does not require a separate processing

[2]See G. Ilie and W. Thompson (2006), (2011)

mechanism. Hence, both cognitive and emotional stimuli may be included in one model of emotional perception. The categories in which combined models ultimately surface are labeled as moods, states, or emotions. This line of thought is consistent with neuro-imaging studies by L. Pessoa, which outline that based on empirical evidence, the isolation of brain mechanisms into emotional and cognitive areas cannot be endorsed (Pessoa, 2008). Moreover, in music information retrieval, which also builds the underlying structure of music streaming services, emotion and mood are commonly organized and processed as if they were synonyms.

The second conflicting point in the area of music perception concerns classification versus restriction. The existence of categories is not deniable, for instance in exaggerated emotional expressions. When a stimulus is categorized, certain objects or concepts form groups of equivalent or analogous elements, thereby reducing the complexity of the information coming from the external world. These are easily identifiable since they fit into prototypes that guide the classification of emotional expressions. Different theories of emotion vary in how much emphasis they put on bottom-up or top-down mechanisms that determine "what makes a stimulus emotional, how it is categorized and how it is perceived, with basic emotion theories arguing that it is mainly [...] stimulus driven bottom-up processes" (Brosch et al., 2010, p.390). However, whereas restrictions do not facilitate the process of determining basic emotions, having categories can assist in this matter. Stereotyped stimuli and mental prototypes that include contextual information can make the understanding of basic emotions easier. Humans tend to intuitively partition items and scan them for prototypes, regardless of the overall subject matter. According to J. Cespedes-Guevara, those ideal representations “influence the perception of emotionally expressive stimuli in a top-down manner, creating artificial discrete categories” (Cespedes-Guevara and Eerola, 2018, p.11).

Similarly, P. Juslin challenges the model of discrete categories by arguing that categorical conceptions are creations of the human mind. In his later works, he suggests organizing musical expressions in a multi-layer system.[3] Those can be differentiated by the higher or lower degree of complexity of their coding structure. Thereby, the core layer is constituted by the basic emotions, which can be extended or modified by additional layers of expressions. Those can convey expressions with more intrinsic and associative coding and enable the listener to experience more complex emotions that are more dependent on social context and individual knowledge. Thus, regardless of his basic emotion model, P. Juslin proposes a possibility for assessing music that conveys more complex emotions under certain circumstances (Juslin, 2013). One automated application of such a framework with a multi-layered system has been implemented with Gracenote’s mood taxonomy: it consists of more than 300

[3] See Juslin (2013)

specified mood categories, which are organized hierarchically while being subject to broader mood categories at the top level. This metadata on moods can directly be derived when using Gracenote's proprietary content analysis and machine learning technologies, without manual tagging. However, in this case, the recognized music emotions go far beyond the standard emotions and moods, occupying terminology with contextual connotations. Thus, descriptors such as romantic, sentimental, fiery, or easygoing are attributed to granular layers.

The third problem in the field of music perception is the question of default versus varying expressions. Numerous studies suggest that recognition of emotions in music depends on multiple perceptual mechanisms. While features such as tempo and loudness enable the detection of different levels of arousal, differentiating discrete emotions depends on acquired knowledge. "In sum, contrary to the predictions of Basic Emotion theory, perception of the whole set of basic emotions in music does not occur early in development, and it seems to depend on learning culture-specific cues such as specific associations between mode and mood" (Cespedes-Guevara and Eerola, 2018, p.6). On this basis, people from different backgrounds and cultures react differently to the same music, which is why emotions elicited by music should be considered a highly subjective phenomenon (Nawaz et al., 2018). For a long time, it was argued that common emotion concepts were innate to us since they root in common discrete biological substrates. However, though intense emotions involve changes in facial and vocal behaviors, not every type of emotion is associated with a distinctive pattern of physiological and expressive behaviors. This issue occurs within not only the basic but also the circumplex approach. Thus, despite its inclusion of arousal and valence, the latter seems to be too reductionist because "two emotions that are placed in the same position in the circular matrix may be very different" (Juslin, 2013, p.4). For instance, this can occur with anger and fear: they are placed in the same quadrant due to their similarly high values in arousal and unpleasantness, but their expressions are not always congruent.

Thus, the assessment of emotion needs to go beyond the categories of arousal and valence to prevent ambiguity and determine contextual conditions under which these dimensions become more salient. To this end, discrete meanings are combined with the listener's top-down knowledge from "past musical experiences, information about his or her current affective state, and cues about the meaning of the event where the music is playing" (Cespedes-Guevara and Eerola, 2018, p.13). Those top-down mechanisms that consider the interaction of a stimulus and the needs, goals, and knowledge of the observer have been shown to improve performance on categorization tasks. This leads to the conclusion that emotional categories are not universally innate but shaped by cultural top-down factors. Furthermore, contextual or cultural information has proved to significantly influence the outcome of categorization (Brosch et al., 2010). Another supporting factor for this approach is that music communicates fluctuations of affect, which can be mapped onto many possible meanings via associative mechanisms. Listeners can experience a variety of emotional percepts depending on the characteristics of their personal, situational, and cultural context. P. Knees and M. Schedl outline a strong influence of the user's context on similarity perception and preference. Thus, listeners can accept different degrees of variability in different scenarios. For example, when exercising, one might accept some musical attributes that one would always skip while studying (Barrett, 2006). Thus, emotional categories should be perceived as adaptive and flexible compilations of emotions.

In sum, all the concerns discussed above indicate that a newly revised analytical approach is required to overcome conflicting aspects between new research claims and partially obsolete arguments. A constructionist account of the perception of musical emotions could provide a solution – one that does not apply basic emotions nor neglects listening contexts. While implementing those main targets, the constructionist account highlights the importance of considering more emotional dimensions and the significance of contextual variance (Cespedes-Guevara and Eerola, 2018). While the conflicting claims within this section have outlined the need for a revised approach, the following section consolidates the relevancy of contextual input as a central element for new frameworks concerning the psychological perception of musical content.

2.3.2 Indicators for Contextual States

Despite the widespread assumption that musical expressivity is organized around a limited set of discrete, biologically predetermined basic emotions, serious theoretical and empirical arguments contradict this, as shown in Section 2.3.1. This is affirmed by J. Cespedes-Guevara: "Evolution has favored flexibility over rigidness, and the communication of social intentions over emotional states. Emotional expressions vary according to the characteristics of the situation" (Cespedes-Guevara and Eerola, 2018, p.12). Thus, a shift from basic emotion theory towards one that encompasses more flexible psychological states and contextual relations should resolve past conflicting findings. To pave the way for this new approach, the research focus must shift from the identification of associations between musical structures and emotion percepts, to the identification of the circumstances. There are both emotional and non-emotional states. This raises the need to incorporate more measurement metrics relevant to this matter in order to model well-suited acoustic experiences for all emotional dimensions and situations, as outlined by the conceptualized framework in Section 3.2.2. Thus, the following highlights the relevancy of contextualized processes in music streaming, before discussing the indicators and operational processes for detecting different states.

Psychological studies on contextual aspects in music perception demonstrate the relevancy of seeing our emotional response to music in context, by emphasizing that this relies on a synergy between musical, personal, and situational factors. This assumes that emotional or musical meanings are not inherent in a sound, but rather emerge from the interaction of the knowledge and aspirations of the listener and the characteristics of the situation (Cespedes-Guevara and Eerola, 2018). One focus of music psychology research is on musical preferences, where German-speaking music psychologists have done significant pioneering work. Preferences are not innate, but acquired by experience and value orientation. Consequently, they reflect, to a degree, the value and expectations of the socio-cultural environment,[4], as they are partially predetermined by them. While the relationship between music preferences and life attitudes is only marginally detectable, musical preferences and situational conditions show significantly stronger correlations with each other (Kloppenburg, 2009). In addition, personality factors in contextual settings become more important, especially in negatively perceived situations, as revealed by Behne and Gembris' studies.[5] Hence, the remodeling proposed in this study aims to assess experiences of anger, sadness, fear, and others without tying their phenomenological character to stereotyped, specific patterns of somatovisceral activity, brain activation, or behavior, and instead offering a wide

[4]The connection of musical and social structures is claimed by Adorno, but not empirically verified (de la Motte-Haber and Neuhoff, 2007).

[5]See Klaus-Ernst Behne (1993) and Heiner Gembris (1990, 1994, 1999).

variety of musical-intrinsic characteristics. The context as well as musical and emotional knowledge are taken into account when constructing the meaning of musical elements. For instance, when the musical materials differ from prototypical stimuli, it can be revealing to examine an individual's range of emotional experience by assessing individual variations in emotional granularity or cultural discrepancies on emotional experiences (Barrett, 2006).[6]

This supports the idea of creating a less rigid typology model that turns away from criteria of distinction such as age, social class, gender, or location and towards a framework closer to contextually relevant subdivisions. However, when considering other options, one realizes that intellectual, physical, and emotional dimensions are very diverse and difficult to compare, which makes it challenging to establish evaluation criteria. There are no longer clear hierarchies for aesthetic values that could explain a universal understanding of art. Accordingly, audiences can no longer be typologized according to the degree of their understanding of art or social classes, but must instead be typologized according a category, such as activity context. Therefore, a newly revised typology needs to be organized around functional niches to which the act of music listening should cater. Today, music oftentimes accompanies non-musical activity and is chosen to reinforce this particular activity in some way by "affecting a psychological state which impacts on desired outcomes. In these contexts the music may not be the primary focus of attention or concern – the focus is rather on its effects" (Sloboda, 2012, p.437). With this statement, J. Sloboda points out that surrounding contexts shape the purpose of the listening process. Thus, the listening takes place not for the purpose of listening to music, but rather to create an atmosphere, or to optimize or change a state. Furthermore, two inherently different practices of music consumption emphasize once more the influence of situational context and intrinsic motivation. Those are, on the one hand, music as an everyday, lifestyle-forming element in the background, and on the other hand, music in the center as an extraordinary event. Those two scenarios are additional to seeing music as a means of identity formation as well as a mood management tool (Smudits, 2007). This is all integrated in the lifestyle that interprets symbols of self-expression in a social context, embedded in the emotional system of a human being, characterized by goals, attitudes, and value systems, as well as feeling-inducing effects of the environment (de la Motte-Haber and Neuhoff, 2007). The interplay of those factors ensures that an approximation to the social reality is pursued.

In the past 30 years, research on the effects of music in commercial or leisure settings has increased due to a growing interest by psychologists in socio-psychological motivators, as well as a rising interest by consumer psychologists. Nevertheless, it has to be noted that

[6]This biasing effect has also been tested in the field of music, when research participants rated music with non-emotional terms such as sharpness, weight or temperature (Cespedes-Guevara and Eerola, 2018).

while musical meaning is established by contextual and situational elements, music-intrinsic features should not be overlooked, since they can potentially have a high impact on those elements. Studies on what music conveys can be helpful in this regard, by analyzing musical parameters for their relation to music-external associations. For instance, musical dimensions, such as rhythm, harmony, or timbre, can help in catering to different audiences when they are conceptualized as music-intrinsic features with distinct means of expression (Herzog et al., 2017). This area of research has established the use of psychometric instruments, such as the General Music Branding Inventory (GMBI), which allow the assessment of associations of attributes induced by musical stimuli. The GMBI's results uncover associative semantic meanings between attributes, such as young, urban, playful, or trustworthy. Hence, music branding can be understood as a tool for sign-based communication, whereby the signs are controlled by high-level (music-related) and low-level (sound-related) audio features, among other elements. It is therefore essential to consider the listening context in order to assess which of those two features should be emphasized. This decision is determined by the awareness and state of a listener, which is in turn mainly determined by preferences as well as situational factors. These ultimately establish whether high- or low-level audio features are best received by a certain audience segment.

Furthermore, audio features can be consulted to strategically create content as per desired expression. This is possible based on tests that have uncovered that low-level features are crucial in the prediction of both arousal and valence. In addition, rhythm features are important for arousal detection, and tonal features greatly assist in detecting valence (Grekow, 2018). Along with those use cases within the scope of contextual listening, industries that make use of audio branding have shown increasing interest in tailor-made audio profiles, among others. Egon Brunswik's lens model describes the communication process of music branding, starting with a brand identity and ultimately leading to a brand image by means of multiple musical metrics (Herzog et al., 2017).This enables one to address listeners in a more targeted manner and implies a gain in attention. For example, expression- or topic- specific adaptations to playlist rotations allow for dynamic structures. This makes it possible to take into account the time of day and activities that shape everyday situations with their specific dynamics. Consequently, contextual labels allow a denomination of categories with commonplace attributes that the majority of people can recall even without musical knowledge. This is provided by titles referring to the time of day, weekday, mood, activity, or social setting, such as Morning Motivation, Workout Beats, or Deep Focus. Moreover, arranging related pieces by context or genre helps the listener to overcome paralysis when inundated with content. Algorithmic tools that are built based on this knowledge assist proactive listeners in navigating through the vast digital music archives and discovering new titles within the restrictions of the chosen context. Furthermore, according to M. Herzog, such algorithmic recommendations allow music to fully unfold its "functionality of mood-management, social-bonding and distinction, identity formation or any other kind of ritual affect-laden everyday use" (Herzog et al., 2017, p.1).

As outlined above, the affective percept is processed by associative mechanisms that "integrate information from past knowledge, contextual information, and the listener's current psychological state" (Cespedes-Guevara and Eerola, 2018, p.14). The methodological choices proposed in this section present options for researchers and creators alike to move beyond the basic emotion paradigm and analyze experiences in relation to the context in which the listening process occurs. Furthermore, the derived learnings can assist in building bridges between psychology and other disciplines that are interested in understanding people's perception of music experiences.

2.4 Progressive Features of Music Streaming Audiences

The aspects regarding consumption, behavior, and psychology that have been outlined above help us understand what, how, and why individuals engage in the act of listening. This scientific analysis incorporating a multitude of relevant disciplines is necessary to comprehend how user needs, interaction behavior, and psychological perception have been affected by the shift of industry fundamentals. Studies have shown that a shift from analog to digital products is feasible once all general needs of media products are fulfilled. However, this does not immediately pose an intrinsic value to listeners. An audience first needs to be apprehended to allow a supply based on its specific preferences.

Firstly, increased listener value can be created by incorporating new content elements. These belong to the context model, which is part of the distribution model. It moves away from the access model by offering increasingly autonomous functioning and recommendation-enhanced elements, in addition to control-related features that rely heavily on a user's input. Illustrative elements and descriptive elements open another dimension, while interactive elements allow users to act upon their urge to function as prosumers. In addition, elements that deepen the artistic concept around the audio product can be consolidated.

Secondly, the observation of listener types in terms of behavior has shown that similar degrees of readiness and activation are present in lean-back as well as lean-forward types. Even individual thresholds are not clearly definable per user type, but instead depend highly on the strength of incoming stimuli. In the last years, there has been a rapid rise in listening being performed as an accompanying activity. During such activities, the neuronal and motoric activation pattern does not lower among listeners, despite the lower interaction rates. To increase the value proposition for these users, context is a major attribute and is being incorporated by more recent technological additions. Overall, a tendency towards less interaction across user types is recognizable, which concurrently entails an increased demand for personalization. This knowledge about users' interaction potential is essential to identify suitable sources and metrics to ultimately enhance content and product delivery.

Thirdly, the psychological perception highlights users' demand for flexibility. This encompasses the challenge of halting the imposition of artificial structures and restrictions upon organic processes. Thus, moods should be treated in the same way as emotions, classification should be a guiding sketch rather than an enforcing limitation, and lastly, we should recognize variations in expressions instead of expecting default patterns. Once acknowledged, contextual interrelations become evident and increasingly dominate on-demand listening. However, those open and flexible structures of emotional perception demand complex evaluation to lay out a system that enables the detection of dispositions, moods, and emotions in an automated manner.

In theory, all of the above aspects within the three given categories enable music creators and curators to create more value before, during, and after the listening process. This empirical research generates a substantiated foundation that displays major connecting points. These may be consulted to solve current problems or presently suboptimal configurations in the music streaming landscape. Based on the derived insights on progressive characteristics of music streaming audiences, the following chapter focuses on finding solutions to enhance the understanding of dissected groups of listeners by consulting an individualistic process of concatenating data analytics.

Chapter 3

Analytical Partition of Listeners

The scientific insights in Chapter 2 uncovered aspects that require renewed techniques to categorize and comprehend audiences in a manner appropriate to current habits in music streaming. The investigations outlined instances in which listeners adopted new manners based on changes in music consumption, listening behavior and psychological processes connected to music streaming. These research findings highlight the importance of varying states and listening contexts. Those form the basis for approaching the stated research question of whether a newly revised audience segmentation method could allow for a more contemporary and pertinent understanding of universal listener demands. The derived audience segmentation in the following chapter serves to uncover differences between listeners while establishing a typology with enhanced methodologies substantiated by data analytical insights.

3.1 Methodology

Originating in the highly individual listening preferences and intrinsic behavior of music streaming audiences as proven in Chapter 2, the deficiencies and shortcomings of existing music analytical tools were discussed. The next challenge was to enhance and combine the processes of those tools to derive new analytical insights. Predominant approaches were entirely one-dimensional, which means sales rather than consumer oriented. This focus on quantitative sales metrics and the failure to address the consumer side resulted in skewed analytical results. Thus, the applied methods sought to create a better understanding of listeners and their behavior and needs by equilibrating metrics of a user-driven with a typically sales-driven perspective.

An empirical research approach was used to test the theoretical hypothesis by extracting information using quantified methods. Grounded theories built the basis for this, and Chapter 2 reveals the background and newly derived insights in this respect. In the manner of an inductive method, the process originated in singular cases in the segments of consumption, behavior and psychology and leads to an overall theory. Following the establishment of overarching principles, a prototypical framework for audience segmentation in music streaming was composed, as circumstantiated in Section 3.2.2. The incorporation of the listening context and the listener's motivation served as key drivers to attain this typology. Thus, traits of perception and interaction with the medium of music are interwoven, while typologies based on psychological studies standardly refer to a specific classification of phenomena into categories based upon typical traits of a personality. This listener segmentation provided the fertile ground for the ultimate analytical method, a uniquely derived multi-dimensional approach that served to fulfil the main goal of combining sales-centric with user-centric parameters. Those theories were refined as more data were sought and were ultimately enriched by case studies.

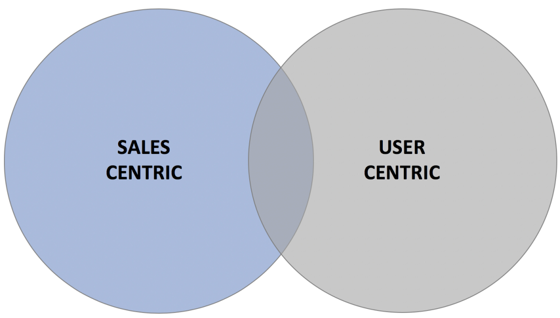

Fig. 3.1 Combined Sales- & User-Centric View

The multi-dimensional approach was based on concatenating data analytics and roots at the intersection of consumer research and business analytics. In this context, the term multi-dimensional indicates the variety of sources, comprising streaming platforms (open-source and content-owner), market research data and benchmarks of behavioral science. Concatenating data analytics describing the alignment and nominal and ordinal consumption metrics attained from a purpose-built database originating in the Spotify's application program interface (API)[1] and market benchmarks allowed for quantitative assessments. Behavioral metrics and audio features represent in-between data, which first require a conversion to be comparable with a common denominator. Last, performed activities, emotions or prior musical knowledge deliver categorical insights that are quantifiable into tiers on the basis of natural language processing. This way it was possible to combine insights from the fields of behavioral science, psychology and market analytics as well as annual industry statements, Spotify's application programming interface and a purpose-built token database. The technical foundation to attain those insights was established by data retrievals. Python queries based on user tokens were linked to a private webhost. Its consumption- and behavior-centric output was further enriched with universally available audio features per track that were grounded on Spotify's tool *Echonest* as well as natural language processing for text-based dissections. The main data source was rooted in the Spotify API and allowed for tracking playlist content as well as accumulated, anonymized user codes called tokens, which are analysed based on four sets of features afterward (see Figure 3.2).

The research medium had to be a suitable music format that would allow for extraction of as many metrics as possible that display a relation to context, listener and listening behavior. Since tracks that are released alone-standing are most frequently consumed without overarching listening contexts or even detached from an album, that medium could not adequately serve as the data source. However, the rise of music streaming over the last few years positioned playlists at the center of music consumption, covering more than half of all accumulated streams in 2018 (Fuller, 2018). Choosing playlists over tracks enabled the collection of information from more users and products within the same timeframe, because playlists allow for capturing all required insights with viewer shots with more interconnections. The minimum number of tokens that contribute to an outcome is set to cover at least 100 anonymized IDs per playlist per week. For this representative sample, groups as per gender, age and territory were extracted per playlist.[2] The aggregation process

[1]Spotify was chosen as the most suitable platform for this investigation because it offers the most extensive amount of open source data and interaction details. However, all results of this framework can be applied to content from other music streaming platforms without reforms.

[2]The data retrieval includes all Spotify territories as per January 2019, - excluding Asia, India, Russia and parts of Africa - focusing on Europe and America.

has been conducted for 10,000 playlists, incorporating all musical genres, for a timeframe of 4 months from January 2019 to April 2019. Data on interaction parameters and audience demographics, as encountered in the case studies, were provided by Spinnin' Records. The label's playlists and playlisted content have been probed for Skips, Saves, Discovery, Completion, Location, Age and Gender.

After querying the database on a weekly basis, the extracted raw data were cleaned, extrapolated and combined. This allowed for pattern detection or the derivation of insights through clustering methods, correlation analysis and indexation. All this contributed to enhancing current methodologies and investigation techniques. Subsequent to the evaluation of the underlying data, several clustering methods were conducted on the dataset. Some challenges emerged, because the clusters of the hypothetical audience segments overlapped to a large extent and were thus not clearly delimitable by this method when applying distribution-based clustering with a Gaussian mixture model. Correlations had to be observed of more than two variables at once to derive the ultimate indices. Another challenge posed the encounter of mixed variables, containing both continuous and categorical variables. Possible options to overcome this concern were given with Gower's distance metric, paired with supplementary techniques.

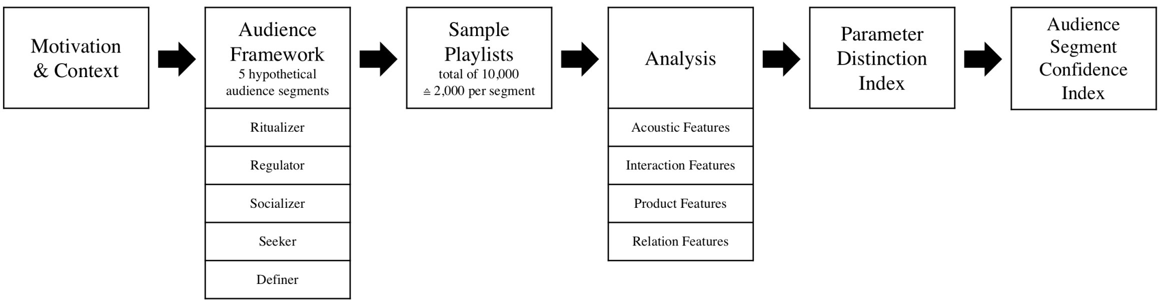

Fig. 3.2 Methodological Approach

3.2 Junction of Analytical Methods

3.2.1 Evaluation of Analytical Models

Prevailing approaches and tools for listener segmentation in the subject area of music branding and streaming primarily take sales-figure and sales-oriented consumption metrics into account. Those can readily be retrieved and analyzed because they always exist in a quantified state. Simultaneously, insights to the user as a persona with his behavior and aspirations are being sacrificed. Because of this negligence of behavioral and personal data points and preference of sales metrics, this approach is one-dimensional. Within this one-dimensional approach, media sales and consumption are used for the analysis of the relationship between media and their audiences, mostly measured in viewing amounts and revenue. Currently available tools underlying this one-dimensional approach are being offered by multiple independent providers. One relevant player in music data aggregation is Music Connect, developed by Nielsen Holdings. A second is The Next Big Sound, which is linked to the streaming service Pandora. Both perform music analysis in four verticals, namely sales, airplay, streaming and selected socials. Although the first aims to provide overviews of market performance trends, the second displays such trends with an emphasis on an artist's popularity. However, more situational and user-specific insights are required to obtain deeper insights that the one-dimensional approach cannot provide. The uncovered weaknesses of the one-dimensional approach include, first, that quantitative metrics are preferred over qualitative metrics; second, that sales-oriented consumption metrics are preferred over behavioral, psychological and geographic data; and third, that no correlation with behavioral or contextual insights is possible.

The following outlines how the multi-dimensional approach is engineered to overcome those deficiencies. When scanning the interlaced metrics included in the one-dimensional strategies, it becomes obvious that one-dimensional strategies solely include quantitative metrics. At present, qualitative and quantitative research methods are only rarely combined. However, assessing the value of each of both research methods helps to identify the assets of each. On the one hand, there are quantitative methods that put an emphasis on objectively measured results and the statistical, mathematical or numerical analysis of data. It focuses on compiling numerical data and "generalizing it across groups of people or to explain a particular phenomenon" (University of Southern California, 2018b). The main goal of quantitative research is to classify features by deriving statistical models and its relations. This aims to explain the facts of the case while highlighting patterns in a measurable way. On the other hand, qualitative research methods imply a focus on qualities of individuals, processes and meanings that are explored in an exploratory and interrogative manner. It helps to understand the attitudes and mindsets of participants. A qualitative analysis displays an orientation toward unique cases and is context sensitive, which is, as mentioned in Chapter 2, an increasingly relevant aspect. In this manner, "qualitative research can be used to vividly demonstrate phenomena or to conduct cross-case comparisons and analysis of individuals or groups" (University of Southern California, 2018a). The focus is on complex interdependencies of qualitative parameters that can include questions about occupation, attitudes, values, lifestyle, knowledge, benefits, consumption habits and usage occasions. Despite the limitations of qualitative research methods, such as non-replicability, they can deliver valuable insights that can assist in interpreting the gained quantitative results. Further, they can facilitate the validation of prototypes and measurements of improvements. To that end, the combination of qualitative and quantitative data can ensure that limitations of one type of data are balanced by the strengths of another. The purpose of combining data is to enrich, examine and explain the results. This leads to a triangulation of confirming, reinforcing and rejecting of results. On the basis of the listed benefits, the multi-dimensional listener segmentation presupposes the inclusion of somewhat-qualitative parameters.

However, because qualitative methodologies would not allow for extraction of insights of large sample sizes, another option had to be found. This is an imperative, because the derived approach aims to derive an overarching audience framework based on superior, context-related attributes that is elevated above demographic and genre limitations. The challenge was in finding an option to best aggregate and concatenate data points that refer to behavior, interaction, sentiments, knowledge and activity in large amounts. Those can all be derived from the API of the respective music streaming platform. The later-derived framework serves as the qualitative instance for the verification process, based on the motivation and prototypical contexts for each audience segment.

To guarantee the option for numeric cross-referencing, two data types assist: categorical variables which, are also known as discrete or qualitative variables, and continuous variables, which are also known as quantitative variables. Continuous variables can be further categorized as either interval or ratio variables. Hence, some of those insights have been transformed into numerical states and others were conceptualized by concatenating multiple measured behavioral and consumption metrics. To derive distinguishable and reproducible statements from those variables, it is necessary to align all continuous and categorical parameters. In the beginning, Gower's distance metric seemed to offer a solution to facilitate the combination of numeric and categorical data. However, further investigations revealed that although the metric "displays the potency to reach high fitness, it could possibly hold a tendency for biases, which could be enforced by the chosen weights and data"(van den Hoven, 2015, p.2). As such, transformation of the data types was conducted in two slightly diverging manners to attain a pure data set. First, variables were established to measure the distance between different parameters or instances. Second, nominal variables were transformed into continuous variables to replace the current variable with N binary variables. In this case, N is the number of values the nominal variable could be. Those methods allow for generation of aligned data formats, which are needed to perform clustering and further statistical analysis.

The selection of the metrics was made more tangible by categorizing those into three superordinate streams of data. The three overarching categories shaping the analytical procedure were context, consumption and sales data. The section of consumption can further be broken down by underlying data sets, which were of behavioral, acoustic and demographic nature. Subordinate metrics can be further expanded as required for particular cases, as exemplified in Section 3.4. In this manner, behavioral consumption data can enhance numerical consumption data. Thus, instead of only focusing on numerical sales and consumption figures, the results provide a more holistic view about the user, his behavior and the consumption cycle. This is ultimately due to a focus on consumer- over sales-centric parameters, which in its last instance may be compared to characteristics of different audience segments in a verification-like process, which is established in the following.

Many underlying data points are needed to achieve personalized or categorized targeting of music. The insights on the drawbacks of current methods in music analytics help disclose new ways to drive understanding of music listeners. This would in the end enable determination of not only what and how much music is being consumed but also by whom, when and how. This could eventually enable the automatic creation of personalized user experiences tailored to one's listening type to attain higher acquisition and retention rates. For this actions, needs and behavioral patterns of the users are increasingly analyzed and combined within multi-layered approaches. Different metrics may be called on depending on the demands of a query and the focus of the musical content that is to be created. One will only hardly ever encounter the need to pair all mentioned metrics simultaneously. However, the outline of the full spectrum is necessary because the repetitive process of data acquisition needs to be laid out in such a manner that the available datasets are accessible from all metric levels. The data infrastructure needs to be set up holistically, offering options to gain insights from a user as well as a music content perspective, to guarantee instant access. This is the case if permanent infrastructures for repetitive API calls are desired. However, if only singular queries are demanded, individual data infrastructures can be created that ultimately demand less processing power as well as storage capacity, which makes these kinds of requests feasible for more actors within the music and adjacent markets.

3.2.2 Derivation of Audience Typology

Based on the research findings in regards to music streaming consumption, listener behavior and psychological reception, it becomes evident that for a time-relevant and reproducible partition of audience categories, two factors are required (see Chapter 2). Those are, first, the psychological motivation to listen to music in that specific moment and, second, the interaction behavior that they display while listening to it. Those aspects were constituted as listening context and listening motivation. Those gain in importance because the listening preferences are not primarily determined by demographics or activities, but rather the listener's intention and motivation while performing an activity, which is summarized with the listening context. Thus, this typology is organized around functional niches, relating to contextual states, to which the act of music listening should cater.

On the basis of this two-factor conceptualization, five listening patterns were delimitable. The derived prototypical user segments are the *Ritualizer*, *Regulator*, *Socializer*, *Seeker* and *Definer* (see Figure 3.3). Those ultimately allow the creator or curator to work with pre-defined categories instead of getting lost between the detailed metrics of each data source. Moreover, they position context and consumption before sales in the assessment process, which allows for giving particular importance to those factors that are otherwise often disregarded in industry-related analysis.

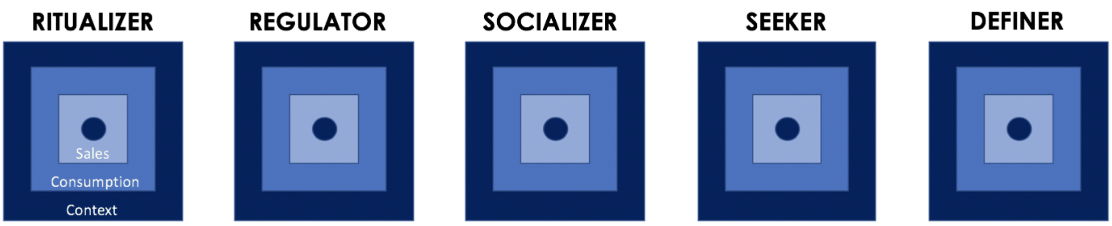

Fig. 3.3 Framework Playlist Audience Segments

The *Ritualizer* aspires to create comfort, order and structure in his day by listening to music in the background. He focuses on low-level audio features while preferring familiarity and longer playlists. He displays a lean-back attitude and makes use of autonomous context content. Exemplary listening contexts are, work commutes, cooking, meditation or morning routines. For these listeners, daytime or an activity or vibe can embody a key entry point for the act of music listening.

The *Regulator* aspires to optimize moods, environments and activities. He focuses on low- and high-level audio features while being highly focused and disliking sudden changes or distractions. He displays a lean-back attitude and makes use of autonomous context content. However, he simultaneously has an above-average interaction potential. Exemplary listening contexts are studying, workout boosts and motivation boosts. For these listeners, a task, activity or mood can embody a key entry point for the act of listening.

The *Socializer* aspires to gain social credit and influencing others. He does so by listening in groups, often to high-valence tracks within longer playlists. Music listening shifts between a background and foreground activity because of a higher degree of changing outer influences. Exemplary listening contexts are a Friday night out, pre-party celebration, afternoon hangouts, a dinner party or a festival. For these listeners, an artist, genre or vibe can embody a key entry point for the act of listening.

The *Seeker* aspires to search for new highlights, discover, learn or fuel his curiosity. He displays a lean-forward mentality and actively looks for a variety of content and frequent updates to listen to as a foreground activity. While enjoying to more elaborate content, he makes use of all contextual content features. Exemplary listening contexts are exploring any broadly delimited content group, new releases or recommendations. For these listeners, a broad genre, era or hits can embody a key entry point for the act of listening.

The *Definer* aspires to organize, make lists and prune his identity. He does so by having specific expectations while searching for something he has knowledge of. He displays a lean-forward mentality, preferring directive content features where he has full control. He focuses on high-level audio features and perceives music listening unambiguously as a foreground activity, Exemplary listening contexts are searching for specific content or the creation of his own playlists. For these listeners, an artist, specific release or niche genre can embody a key entry point for the act of listening.

Fig. 3.4 Exemplary Playlists per Segment

On the basis of this framework, the five categories can be established as hypothetical prototypes for different pools of listeners and can ultimately serve as a test object for parsing consumption preferences. This facilitates the perception of listenership, contouring who listens, why, how and to what. However, the above given imprecise definitions make it evident that qualitative factors such as motivation and behavior cannot serve as the sole factors needed to group listener profiles. As a result, the chosen mixed-method approach of a multidimensional listener segmentation combined with concatenating data analytics paves the way for the hypothetical typology to be substantiated and tested in a quantitative manner. The objective of the statistical analysis in the following chapter is to define those segments with data excerpts as clearly delimitable categories and prove or disapprove elements adopted from scientific research studies.

3.3 Audience Partition

As per the hypothesis of this research, the following analysis and discussion is subject to the question of whether the establishment of a newly revised audience segmentation method, focusing on consumer- and context-centric data points, can facilitate a better understanding of the audiences' demands by analytical means. Therefore, as stated in the methodology, the aggregate data points aim to substantiate the empirically derived listener typology.

3.3.1 Statistical Analysis

The statistical analysis was undertaken to substantiate the assumptions established with the listener framework in Section 3.2.2. This categorization is based on insights of a multidimensional analysis of 10,000 Spotify playlists as outlined in the methodology section. The main dimensions that deliver impact for the analytical groundwork are product-based and listener-based metrics. Those help to attain distinguishable and reproducible statements for the definition of the derived five categories and therefore determine the structure of the subsequent analysis.

Base category	Property	Parameters
Product	Genre properties	top track genres, artist genre mix
	Cycle properties	release age track renewal cycle
	Popularity properties	followers, max. artist monthly listeners, avg. track popularity score
	Acoustic properties	energy, danceability, acousticness, valence, BPM, liveness
Listener	Relation indicators	activation, emotion, knowledge
	Interaction indicators	saves, skips, discovery, completion

Table 3.1 Categories, Properties and Parameters of Analysis

Product-based metrics

Track Genres are assigned to a playlist when a genre captures 10% of the playlisted tracks. When observing the top genres, as per genre tags of the Spotify algorithm, one can identify that of all playlists, half are built on 3 genres and the other half on 4 genres. Thus, the mean lies between 3 and 4 genres. This indicates that few and many genres are included in playlists at about the same rate. However, the extremes display a slight shift, demonstrating a higher occurrence of playlists with fewer genres: 30% of all playlists have 2 genres, but only 8% have >6 genres (including up to 10 genres). For instance, 3 genre playlists appear most often in the Ritualizer and Definer segments. The development across segments is best explained with a u-shaped distribution from the left to the right segments. A genre count of 4 occurs most often in the Socializer segment. This is followed by the Regulator and Seeker playlists, which induces a bell-shaped distribution across the axis of the segments, indicating a greater distribution and accompanying degree of flexibility than in the other segments. With Ritualizer and Definer over-performing with the lowest genre count and Socializer over-performing with the highest genre counts, the extreme examples also verify the stated focal points.

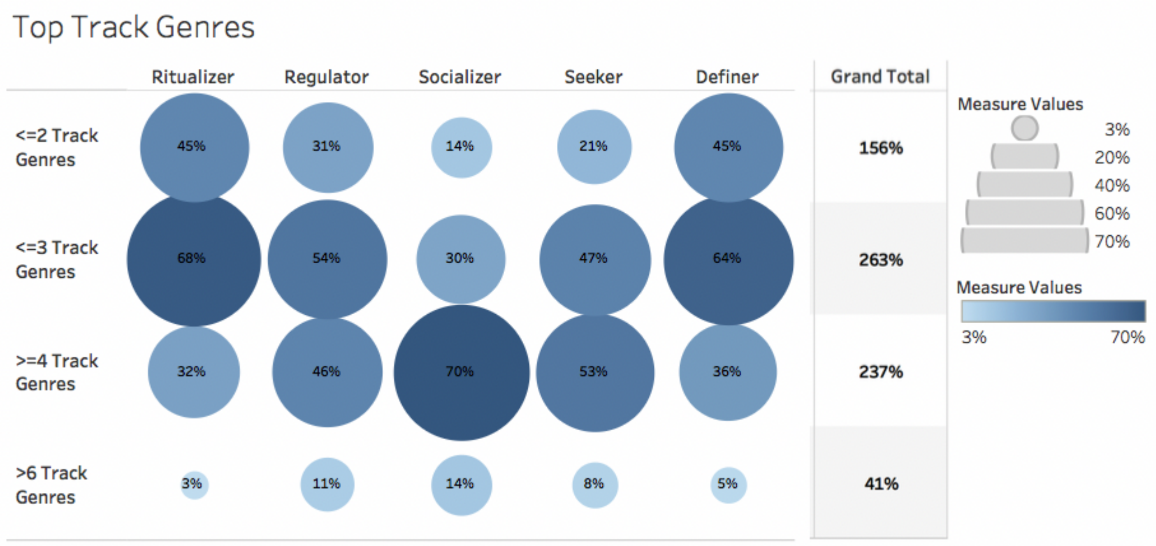

Fig. 3.5 Analysis: Top Track Genres

Artist Genres are assigned to the artists of a playlist and indicate the main sound sphere of the playlisted artists. Those top genres of the artists allow for communicating the sound of the playlist based on the artist's overarching sonic reputation and sentiment, rather than only one-track tags, as in the prior example. Playlists with clear genre directions within its primary genre are most often found in the Seeker segment, closely followed by the Definer and

Ritualizer segments. Socializers heavily underperform within those given metric limitations. Playlists in which the artists brought in two highly relevant genres occurs most often in the Definer followed by the Socializer segment. Ritualizers underperform regarding the high relevancy of two genres. The average share the top two genres cover is the highest in the Definer and the Seeker segments. The Socializer segment's preferences are visible in the last tab, which states that the top two genres on average only cover about 35% of the genre spectrum offered by its artists. This ties into the insights from the track-level analysis, which demonstrate that 70% of all Socializer playlists contain 4 genres in their tracklists, leaving capacity for a variety of genres in terms of track tags that is reconfirmed by its artists incorporating a great bandwidth of genres.

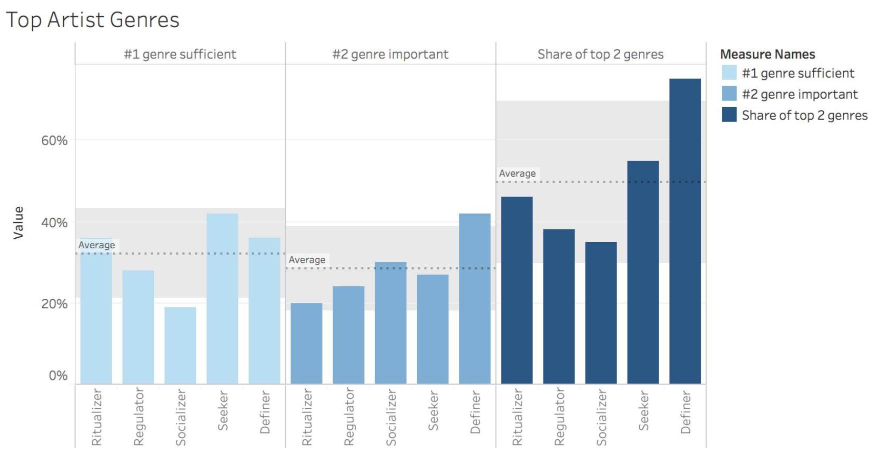

Fig. 3.6 Analysis: Top Artist Genres

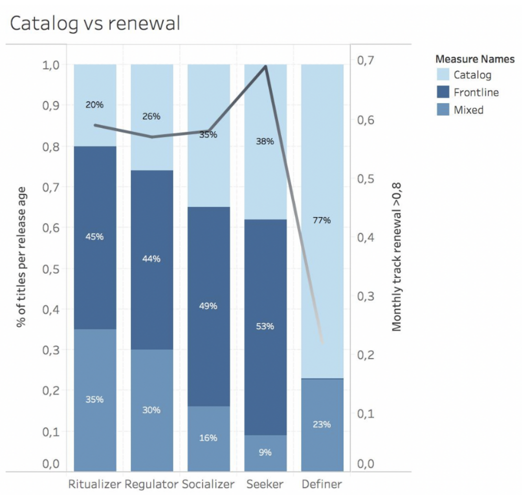

Fig. 3.7 Analysis: Catalog & Renewal

The analysis of the *Track Age* before its addition to the playlist provides further insights on the mixture of the track ages, divided into pure catalog, frontline (genre ratio >65%) or mixed-release-age playlists (genre ratio <65%). A song is considered a frontline up to 18 months after its release, and then it becomes catalog. This mixture of catalog and new releases gives an indication on the intentions, openness to discovery and popularity requirements. In addition, this metric should be considered alongside the popularity, discovery rates and track renewal. When observing the release age of titles within the playlists, it becomes obvious that the main proportion of playlists (44%) comprises mostly frontline content in all segments, except in the Definer segment. Catalog titles as the main contributor increase proportionally from the very left to the very right segment, with the highest difference between the Seeker and Definer segment, an increase of 1.9 times, making up for the missing frontline titles. Likewise, there are playlists that focus on mixed repertoire, with the most on the left end of the spectrum and declining to the right. However, the Definer segment has made this content mix one of its unique additions regardless of the catalog dominating 77% of tracks in this playlist type. The percentage of tracks replaced within a 28 day cycle is mostly between 57 and 70% across all segments, and the lower number is used as a threshold. This discloses that the first three segments have similar renewal rates of >57%. Seekers surpass this score with 69%. All of those segment leave the Definer segment behind, which renews only about 22% of its playlists within 28 days.

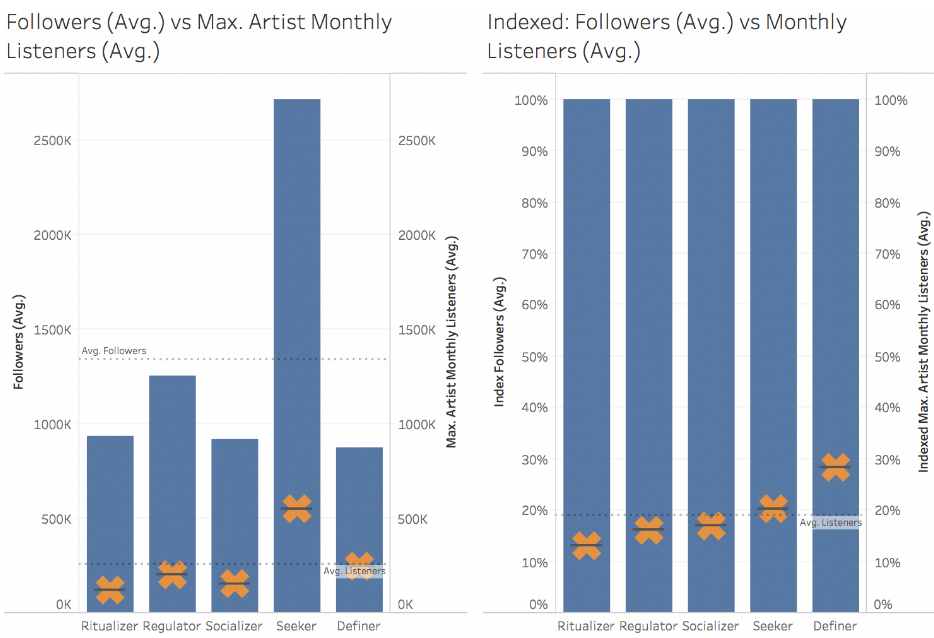

Fig. 3.8 Analysis: Followers & Listeners

When observing the *Average Follower* counts, the data show that Seeker playlists have the highest level of followership and retention rate. However, the monthly listener rate per artist within a playlist allows for putting the follower counts into perspective. Monthly listeners are unique listeners who play the selected tracks/playlists during a 28-day period. Using absolute numbers, Seeker playlists display the highest follower count and potential for monthly listeners, but when indexing the maximum listeners against the playlist followers, a different picture emerges. The indexed context discloses that the highest return of followers per artist can be achieved by playlists within the Definer segment.

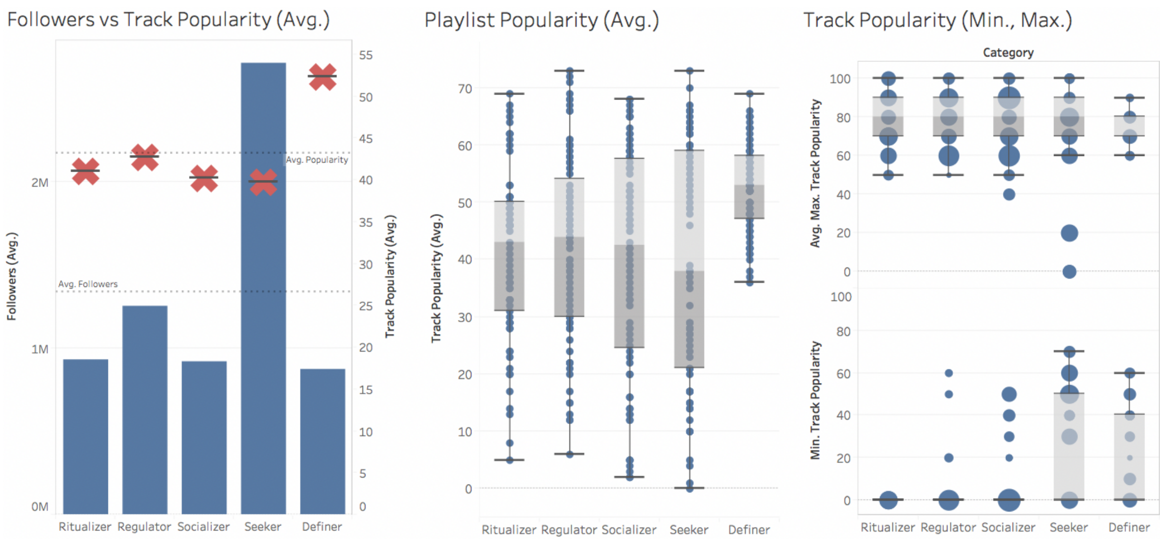

Fig. 3.9 Analysis: Track Popularity

A look into the average *Popularity* score of playlisted tracks within the five segments outlines that only one of the five segments can set itself apart. The Definer segment has an average score of 52 points, with the others located below the overall average of 43 points. When opposing the average popularity scores and playlist followers, it becomes apparent that as the Seeker segment shows, popularity scores and follower counts do not have an immediate relationship.

The distribution of a segment's maximum and minimum popularity scores allows the following insights. Maximum track popularity has its interquartile range mostly between the popularity degrees of 70 and 90, and a median around 80. Socializer and Seeker allow for a wider dispersion of the maximum value, with the Seeker displaying the widest gap. This is due to the discovery of often brand new tracks and new hits by listeners of this segment. Definers are the only segment that is settled a bit lower, with a median also at 70 but the IC limited to 70–80. This is due to the very specific selection and curation of this content, which is in an era, artists or genre context the selection of the most-renowned or -listened-to content.

As for the minimum popularity, as displayed within the Ritualizer segment, the minimum score of each playlist is 0. Thus, Ritualizer playlists are always open to allow for new or unknown tracks.

The remaining segments display clusters of medium and low scores for their minimum popularity scores. Of those, the Seeker segment has the highest disparity and IC followed by the Definer segment. This is due to Seekers yet again looking for brand new content as well as specialized content, which may include the top hits of 1990, which are exclusively of higher popularity. This is why the disparity within the Seeker segment additionally entails a large range and a significant gap between 0 and 30. The Definer segment displays a high continuity and density across all popularity scores up to 60. This is within a large range and is due to a specialization on very popular as well as niche tracks that range across all popularity rates. However, because this user segment directly searches for known content, it is the segment with When observing the popularity in detail by looking at the popularity dispersion within the five segments, it becomes apparent to what side and what extent the popularity of playlists varies. Although the previous graph presented the average popularity of playlists within a segment, the median also presents a similar picture, with only the Definer segment setting itself apart. The narrowest range can be found in the Definer segment, which has its upper and lower hinge at 47–58 and the whiskers at 36–69. This leaves only a 33-point playing field for curation among the most popular content, which is further delimited by the hinges to an 11-point playing field. The dispersion additionally shows that the widest accepted range can be found in the Seeker segment, covering the full range. Regarding the remaining segments, it is noteworthy that the popularity range of the Regulator lies slightly higher than in the other segments, with the lower whisker at 12. Although the median and average are close for the first four, the acceptance of variance for the popularity of tracks is more confined for the Ritualizer, gradually increasing toward the fourth segment. Primarily, when observing the inter-quartile range there is a clear sequence visible starting with the Definer, followed by the Ritualizer, Regulator, Socializer and, last, the Seeker.

Before observing the acoustic metrics per segment based on Spotify's definitions for its audio features, I applied the figures to scales from 0 to 10, with 0 being the lowest and 10 the highest degree of each feature. *Acousticness* indicates whether the track is acoustic, as in solely or primarily uses instruments that produce sound through acoustic means, as opposed to electric or electronic means. Degree 10 represents high confidence the track is acoustic. *Danceability* indicates how suitable a track is for dancing. The calculations are based on a combination of musical features including tempo, rhythm stability, beat strength, and overall regularity. A degree of 0 is least danceable and 10 is most danceable. *Energy* indicates a perceptual measure of intensity and activity. Prototypically, energetic tracks are perceived as fast, loud and noisy. Features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate and general entropy. *Liveness* indicates the presence of an audience in the recording and the connected likelihood of a live performance. A degree higher than 8 provides strong probability that the track is live. *Valence* indicates the musical positiveness conveyed by a track. Tracks with high valence carry a more positive sound character, evoking happy, cheerful and euphoric connotations. In contrast, tracks with low valence carry a more negative character, evoking sad, depressed and angry associations. Tempo describes the overall estimated tempo of a track in beats per minute (*BPM*).

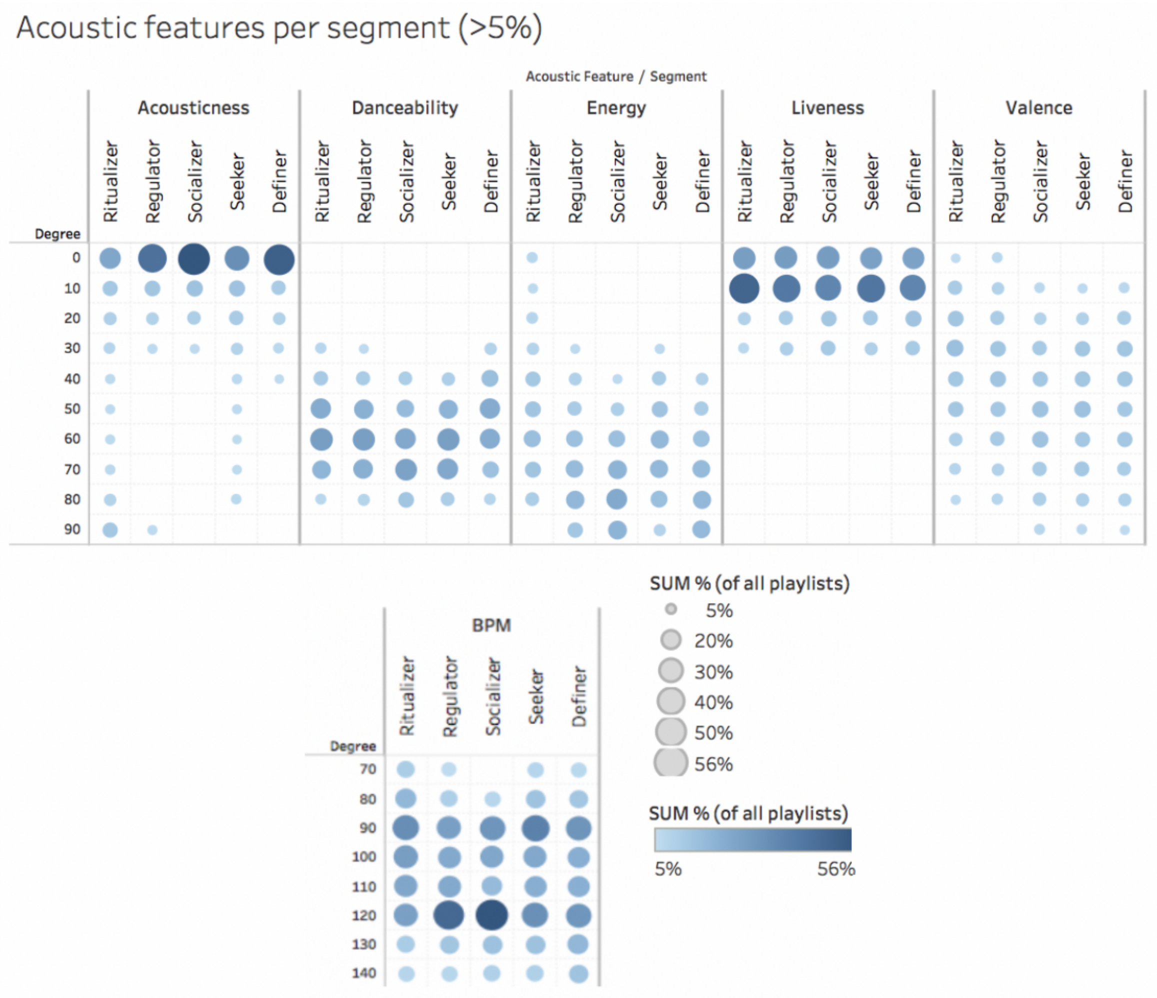

Fig. 3.10 Analysis: Overview - Acoustic Features

When observing the progress of the dispersion of the acoustic metrics onto the range of degrees, one encounters a homogenous picture. Consequentially, the presented data are restricted to data points that capture more than 5% of all listed tracks within a segment to present a more confined view. The focal points are for Acousticness in degree 0, Danceability 50–70, Energy 60–90, Liveness 0–10 and Valence 30–50. Those display the most commonly placed track parameters across all playlists. This surfaces regardless of the playlist's segment. However, when continuing to look into the distribution of the values, a picture of wider acceptance becomes visible in isolated cases. All those acoustic features need to be observed with respect to the normal distribution of values for each feature, which are based on the entirety of tracks within Spotify's catalog. A positive skewed distribution is normal for Liveness and Acousticness with highly discernible peaks. Danceability and Energy are negatively skewed in a normal state, with an overall higher level depicting a bell-shaped curve. Tempo and Valence resemble normal distributions, whereas Valence has an overall higher level than Tempo because of the high starting and end values. A standard distribution for Tempo focuses on the central degrees. The following results can be deduced by comparing these patterns of distribution with the derived distribution curves of the analysis. Acousticness and Liveness correspond with the normal distribution overall. Only the Acousticness feature of the Ritualizer segment diverges slightly. Its peak is not as discernible as of the other segments, and the rest of the curve does not follow a concave up (decreasing) shape but instead predefines a u-shaped distribution. Thus, it is more accepting toward a broader distribution including values above degree 0 with a second focal point on fully acoustic pieces. Source Spotify development sound features Danceability and Energy also correspond with a normal distribution. Only the Ritualizer starts with a slightly higher level of the lower degree of Energy, allowing for the widest distribution of energy levels. Valence yet again corresponds with the normal distribution, only varying slightly with a positively skewed distribution for the Ritualizer segment toward a more negative sentiment and the Socializer segment toward a more positive sentiment. Tempo displays a clear peak on degree 120 in the Regulator and Socializer segment and some varying curve patterns. However, combining all segments displays a curve close to the normal distribution and ultimately fits the overall picture. Of all acoustic features, the overall picture when combining all segments resembles the standard distribution as per Spotify's definitions. This verifies a well-selected choice of playlists capable of representing the entirety of the platform's catalog.

Although degree 0 for Acousticness is the most represented degree across all segments, the Ritualizer steps slightly back, with only 25% in the degree 0 slot but continuing its dispersion through the full range. Thus, the Ritualizer segment spans across all degrees with another focus on degree 90. The Regulator segment is likewise open for an expansion in this very remote degree, highlighting an outlier. Simultaneously, the Seeker joins for all segments besides the highest one. Socializer and Definer stay mostly within the lower mainstream degrees. Danceability is overall homogeneous across, though Socializer and Seeker thrive for higher degrees than the rest. This becomes even more unambiguous when looking at the above pictured circle distribution, which includes only values >5%. The Energy feature is predominantly located in the higher degrees, whereas exclusively the Ritualizer segment is open to incorporating tracks from all degrees. Liveness has a uniform allocation, overarching the most homogeneity across all features. Valence displays concurrently a wide dispersion of similar weights, almost covering the full range. One slight shift is worth mentioning. Ritualizer and Regulator have a tendency towards the lower degrees and the remaining for tracks to higher valence scores. Those allocations facilitate a better understanding of the acoustics features of tracks included in the respective playlists. Ultimately, the Regulator segments displays the greatest coverage of degrees across all acoustic features in placed tracks, predominantly because of its specifications in Acousticness and Energy. The tempo (BPM) of the tracks primarily are in the range between degree 70 and 140, exposing that values 0–69 and 141–210 as responsible for a negligible share of tracks.

Although Figure 3.10 gives an overview of the frequency distribution per degree for a certain acoustic parameter, this neglects the distribution within one single playlist. It is necessary to circumvent comparing completely different playlists within one category to each other. Instead it is recommended to promote keeping the overall curation in mind while highlighting outliers. When observing the amount of instances within one single playlist instead of the entirety of levels included in all playlists, the Socializer segment on average makes use of only 8.2 degrees and the Regulator of 8.3 degrees within one playlist. Nevertheless, some playlist might have an overall higher or lower level, which makes the overall BPM range of the Socializer appear as wide as in the other segments. Likewise, for Danceability, those two segments, Socializer and Regulator, have the smallest coverage of values, with the Socializer claiming 4.9 and the Regulator 5.2 instances.

Observing the standard deviation of the population (*STDEV.P*) per acoustic feature makes it discernible that the highest deviation occurs in the metrics Acousticness and BPM with an average standard deviation of 27.1 and 29.1. Danceability and Liveness display the least deviation from the mean. Although the occurring variance among Danceability and Valence are mostly positioned in close distance around the STDEV.P average of its section, among the other sections, outliers need observed more closely. The most divergent metrics are displayed by the feature Acousticness, with the biggest confidence interval because of the high variance in STDEV.P per segment, including two outliers that are located outside of the 0.95 confidence interval. Those are above average for the Ritualizer and below average for Socializer, with a STDEV.P of 33.1 and 20.6. This occurs because of a u-shaped curve with a dual focus on the opposing values of 0 and 90 in the Ritualizer segment, as already outlined. Energy entails one significant outlier, the Socializer segment on the low end (STDEV.P

18.5), with its emphasis on the values 70–90. In contrast, the Ritualizer segment has a widespread even distribution of values, which results in a higher STDEV.P. Liveness also entails one outlier, the Ritualizer segment with (STDEV.P 13.3) 2.0 points less than the group average and not contained within the 0.95 CI. Overall, the audience segments with the highest outlier count are first the Socializer segment with two below average standard deviations in Acousticness and Energy and second the Ritualizer segment with above average in Acousticness and below average in Liveness. Those outliers allow an immediate picture on the overall distribution of the values around the mean. Further Figure 3.11 and 3.12 enables the discernment of whether singular playlists skew the overall distribution in Figure 3.10 or if a uniform dispersion exists among the playlists, within the boundaries of each acoustic feature.

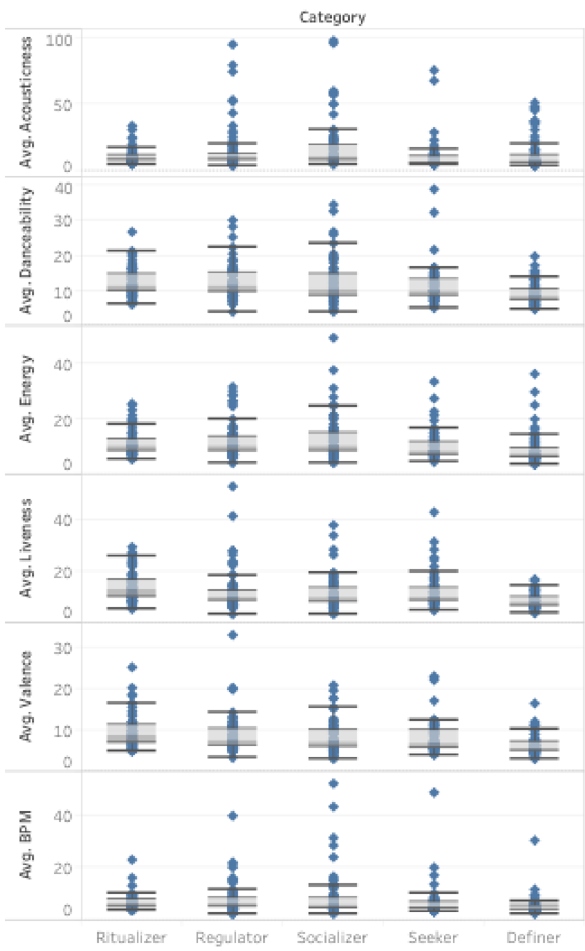

Fig. 3.11 Analysis: IQR - Acoustic Features

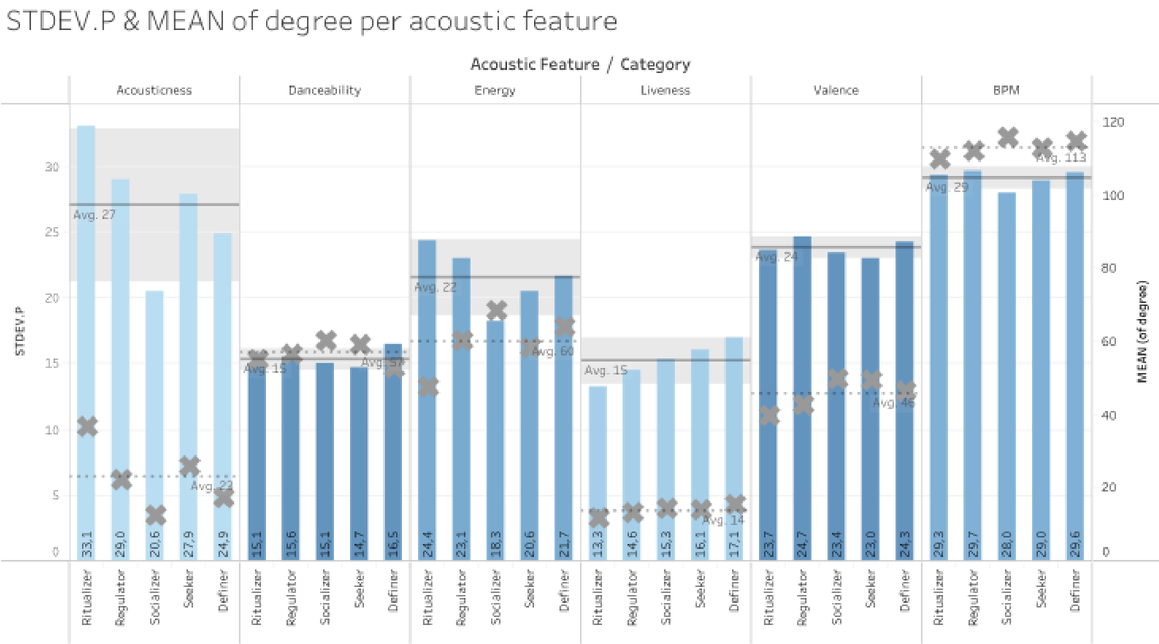

Fig. 3.12 Analysis: STDEV.P & Mean - Acoustic Features

Observing this by audience segment shows that Danceability and Liveness are the two acoustic features with the lowest standard deviation around their means. The two metrics with the highest STDEV.P scores are Acousticness and BPM. Only within the Socializer segment does this order change, where the top positions are BPM and Valence deviation, yet again because of its u-shaped distribution and resulting low deviation in Acousticness. The following insights can be derived by looking at more granular details and observing the correlations of audio features on a per playlist basis. In the Ritualizer segment, there is a high degree of correlation between the four fields of Valence, Energy, Danceability and Acousticness, although they are sensitive to outliers. On the basis of Danceability, the metrics Valence, Energy and Acousticness frequently display correlated movements. In the Regulator persona, there is an interdependency between Valence and Danceability visible, occurring in most cases. Furthermore, on the basis of Liveness and Acousticness, all remaining acoustic features display identical, correlated movements. In the Seeker segment, there is a frequent correlation between Energy and Acousticness discernible. In the Socializer and the Definer segment, there are no clear, reoccurring interdependencies visible.

Listener-based metrics

Property	Listener type	Low	Medium	High
Activity	Ritualizer	47,4%	0,0%	52,6%
	Regulator	54,5%	0,0%	45,5%
	Socializer	50,0%	0,0%	50,0%
	Seeker	100,0%	0,0%	0,0%
	Definer	100,0%	0,0%	0,0%
Emotion	Ritualizer	68,4%	0,0%	31,6%
	Regulator	54,4%	0,0%	45,5%
	Socializer	83,3%	0,0%	16,7%
	Seeker	100,0%	0,0%	0,0%
	Definer	100,0%	0,0%	0,0%
Knowledge	Ritualizer	100,0%	0,0%	0,0%
	Regulator	86,4%	9,1%	4,5%
	Socializer	66,7%	22,2%	11,1%
	Seeker	14,3%	76,2%	9,5%
	Definer	0,0%	10,0%	90,0%

Table 3.2 Measurements of Relation-Indicators

Data on *Relation Indicators* were extracted by means of natural language processing to disclose the dispersion of activity-, emotion- and musical-knowledge-related indicators based on the titles and descriptions of playlists. Those are represented in this research by three properties, *Activity*, *Emotion* and *Musical Knowledge*. The degree of the relation is based on the existence or non-existence of specific terms relating to each property. This definition allowed for the results as outlined below.

First, the two poles of the property *Activity* can be described as follows. High activity relation indicators are based on keywords that describe an activity, action or motion, as found in Power Workout, Dinner Music, Segue o Baile, or Study Zone. Low activity relation indicators do not include any of the mentioned keywords, as found in Hits Unlimited, Coldplay Complete, and Verano Forever. It should be noted that this does not indicate that no activities are associated or factually performed while listening to music. It indicates only that the title nor description of a playlist gave indications that the playlist's main context is connected to an activity.

Second, the two poles of the property *Emotion* can be described as follows. A high emotion relation indicator is based on keywords that describe a state of arousal or calmness or emotion, as found in Sweet Soul Chillout, Happy Hits!, and Confidence Boost. The analysis finds a 2:3:1 ratio among Ritualizer:Regulator:Socializer ratio. Low emotion relation indicators did not include any of the mentioned keywords, as found in Dinner Music, Hits Unlimited, and Women of Pop.

Thirdly, the two poles of the property *Musical Knowledge* can be described as follows. High Knowledge ratings are based on keywords that include festival names, a two-factor combination of, for instance, era and location, era and type of music and artist and genre. Other keywords are genre-specific termini such as Falsetto or Riff, sub-genre categories or localized genre names, or attribution to specific artist. Such examples can be found in Power Ballads, Brahms Piano Concertos, Deutschrap: Die Klassiker, 90s Acoustic and Pop Urbaine. In contrast to the other two categories, Musical Knowledge, allows to define a third segment that lies in between the two extreme poles. Those medium knowledge ratings are based on keywords that include a broad genre, national holiday music, all works, new releases, and top hits of a genre or within a country. Such can be found in Hip Hop Crew, Festa Brasilera, All Out 80s, and Top Hits Portugal. Low knowledge ratings do not include any of the above-mentioned keywords and tend to integrate commonplace words referring to everyday actions or states. Examples can be found in Study Zone, Power Workout, Sad Songs, Girls' Night, and Evening Acoustic. The analysis highlights a large number of low-knowledge playlists. A ratio of 2:1:1 among low:medium:high knowledge playlists outlined a focus on context-related emphasis when searching for playlists.

The insights of the three described relation indicators permit the layout of an exemplary set of playlists for each listener segment. This outlines stereotypical playlists as listened to by the respective audience segment (see figure 3.4).

Property	Listener type	Minimum	Maximum	Average
Discovery	Ritualizer	0,0%	57,7%	41,9%
	Regulator	25,3%	67,0%	44,3%
	Socializer	18,1%	64,6%	45,4%
	Seeker	0,0%	84,8%	44,6%
	Definer	0,0%	65,4%	42,3%
Saves	Ritualizer	0,0%	0,6%	0,2%
	Regulator	0,1%	1,2%	0,3%
	Socializer	0,1%	2,0%	0,4%
	Seeker	0,0%	3,0%	0,5%
	Definer	0,0%	2,7%	0,4%
Skips	Ritualizer	0,0%	5,8%	2,9%
	Regulator	1,9%	10,2%	4,8%
	Socializer	2,5%	14,2%	6,3%
	Seeker	0,0%	17,6%	5,9%
	Definer	0,0%	10,8%	5,2%

Fig. 3.13 Measurements of Interaction-Indicators

Data on *Interaction Indicators* were derived by consulting user tokens, with the main properties being *Discovery, Saves* and *Skips*. The definition of the hereafter analyzed metrics is as follows. A Discovery is defined as the first time a user streams a track within Spotify.

Discovery rate is then the number of discoveries on a given day as a percentage of the total plays. A Save is defined as the first time a user streams a track in their library within Spotify. This also only goes back 100 days in history. A Skip is based on any play that ends between 30 and 60 seconds. Historically, the first 30 seconds is not included, as those plays are not revenue-bearing. This already presupposes that even if details on those indicators are available, it is necessary to evaluate them within a larger listening context.

Commonly, hierarchical clustering of interaction indicators grants the grouping of similar data points. This is used in unsupervised machine learning techniques with the aim of finding similarities between data points and grouping them. However, when comparing the natural dispersion of interaction behavior with algorithmic-generated clusters, it becomes obvious that the five pre-defined clusters overlap too much to be forced into delimited categories (see Figure 3.14). For this reason, the subsequent analysis instead highlights whether an audience segment allows for a large variance or in contrast demand for rigorous restrictions. It then becomes apparent which listener types allow for greater or smaller variance within the given research parameters and how high the group averages are positioned.

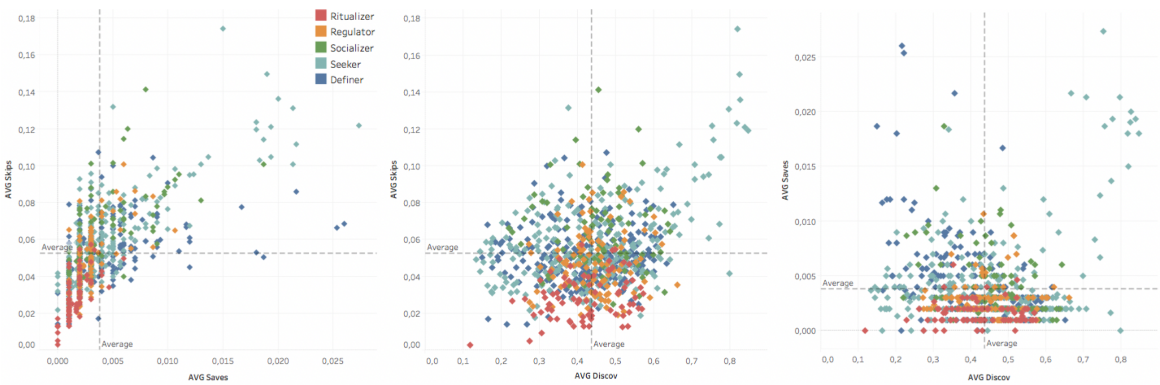

Fig. 3.14 Interaction Scatterplots

Every consumer type interacts differently, which can be shown by a box-plot. All visualizations show the standard variation with the inter-quartile range (IQR) around the median and the whiskers show the range of the first and fourth quartile. Occasional outliers visible as single points are considered to be exceptions within one of the interactive actions despite fitting the picture of their consumer type in regards to other interactions. The IQRs illustrate that the ratios of Saves, Skips and Discoveries have a distinct scales. Discovery rates are centered on 40–46%, whereas Skips rates at 2.88–6.33% and save rates are at 0.16–0.45%. Because the share of all three interactions is always the same across the consumption types, it can be derived that the action threshold for saves is higher than for skips and by far higher than for discoveries. The latter can also be derived from its automated nature. Within the tested medium of playlists, Discovery is mostly programmed or pre-regulated with the choice of playlist, which only calls for a singular action. In addition, the consideration of the percentage average of each group per interaction shows that the distribution of all data points is almost entirely symmetric. There is a normal distribution across all categories. However, when looking at each category per interaction individually, we can uncover some discrepancies.

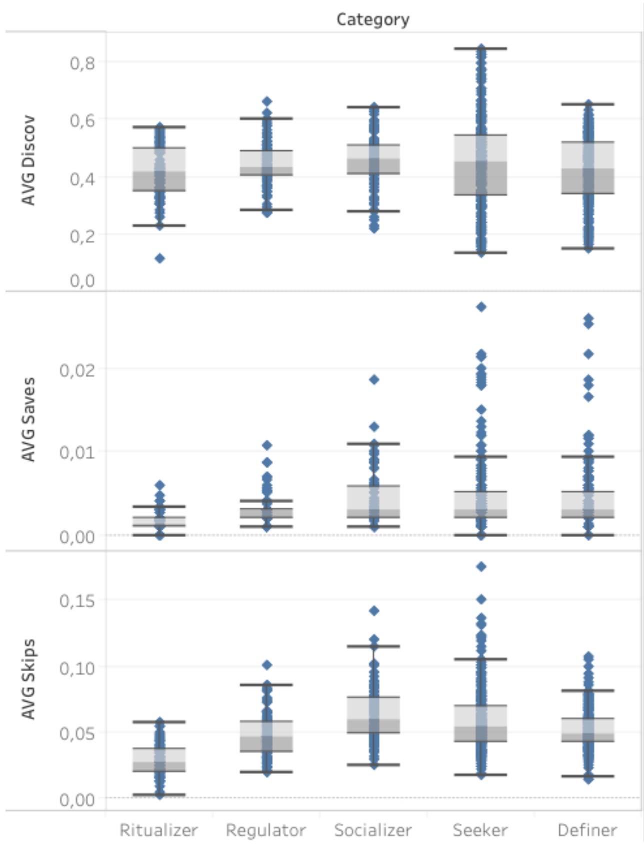

Fig. 3.15 Analysis: IQR - Interaction Features

For the property *Discovery*, the median is located at approximately the same height for all categories. Although the mean is 0.43, the median is located close to it, with a minimum of 0.41 for the Ritualizer and a maximum of 0.46 for the Socializer. The overall median is at 0.4380. The average of Discovery displays the closest relation between median and mean of all three interaction tests. However, for all categories, the mean and median are relatively close, which leads to a non-skewed symmetric distribution. Regulators display the narrowest IQR, whereas the Seeker segment displays the widest one. The Seeker segment further demonstrates a wide range in Q1 and Q4. From the insights of the measured discovery rates, it can be derived that discovery takes place by decision and is not bound to a specific segment. However, the narrow IQR of the Regulator segment indicates that those consumers are more intent in their choice and how much new content they want to be exposed to. Further, there are more personalized playlists linked to listening incentives, typical for Regulators. On the other hand, Seekers are more open to exploring music with a high as well as low tolerance of discovery, depending on their initial choice of the content.

For the property *Saves*, the three segments Regulator, Seeker and Definer display a significant number of outliers. The outliers range far beyond the fence of the whiskers. Because of a very strong Q2, there are left-skewed datasets, with the majority of data on the lower end, in the Socializer, Seeker and Definer segments. Furthermore, we can observe a very narrow IQR, and the Seeker segment displaying the widest one. From the insights of the measured save rates, it can be derived that listeners from the Socializer, Seeker or Definer segments display about the same likelihood of saving a track, with a great flexibility in regards to the maximum.

For the property *Skips*, the median and averages vary the most between the five audience profiles, in contrast to the other properties. The Ritualizer has the lowest median, at 0.027, and the Socializer segment the highest, with 0.063. A significant number of outliers is visible among the Seekers and the Definers. Despite the varying height of the median, the length of the whiskers and the width of the IQR are similar across all consumer types, except the Ritualizer, which displays a slightly narrower range. The insights of the measured skip rates let us derive that Skips are a valuable parameter for differentiation because the segments display unique means and clearly delimitable ranges.

A look at the *Completion Rates* of each segment enables a further categorization. Ritualizers display overall the highest completion. They complete an average of 94.5% of started tracks. This is regardless of whether the playlist is connected to an activity or not, as there is a marginal difference of 0.4% between both. Socializers complete about 88.4%, which is overall the lowest completion. Seekers display a completion rate of 88.9% and Definers 89.3%. This places them before the Socializer segments, yet very distant from the Ritualizers. Regulators complete about 90.7% of started tracks and additionally display a wide difference in completion rate between activity- and non-activity-related playlists. Those can be differentiated clearly, with a 7% gap between a completion rate of 88.8% for the first and 92.7% for the latter. The average completion rate of 90.6% across all categories discloses that the Ritualizer segment sets itself apart from the remaining categories in this interaction format. Simultaneously, this breakdown reveals that the Regulator segment is the one that caters the greatest mix of playlists tailored specifically for activity and non-activity-related scenarios, whereas Seekers and Definers are clearly defined by playlists that do not primarily cater to activity-related scenarios. Another way to break down the Completion Rates is by required music pre-knowledge. This was segmented according to the relation to music specific or mundane labeling and description of playlists as well as the amenability of the musical content by means of natural language processing based on manually selected criteria. This allows for relating back to the proportion of activity-related content in a listener category as well as their initial openness to discovery. This element concludes the results section with the derived analysis of 20 parameters within the scope of five listener segments, as derived within the hypothetical audience typology. However, this section is to be seen only as the first analytical step. An additional concatenation and indexation of the derived findings permitted an assessment of the applicability of the proposed framework. Because the measurements of each unique property are only of added benefit for applicability once they are compiled per audience type.

Concatenation of Insights

The research question demands a concatenation and indexing of the derived analytical results to understand correlations, inter-dependencies and normative patterns per segment. As the analysis of the preceding paragraphs shows, some playlists have a wide range of attributes and others work within very limited restrictions. Those restrictions may vary greatly or may be bound to a certain scope. Thus, it is essential for analytical as well as creational processes to weigh in the *power of differentiation* as well as the actually measured values of product- and listener-based metrics. This creates a holistic overview of the correlation between distinct factors that allow the derivation of prototypical thresholds for each audience segment.

The visualization in Figure 3.16, allows an overview based on a groups average measure value, whereas the power of differentiation, see Figure 3.17, is based on a combination of all statistical validation metrics. To facilitate the understanding of the insights, the circles reflect the measurements in three sizes, with the largest resembling high levels and positive answers.

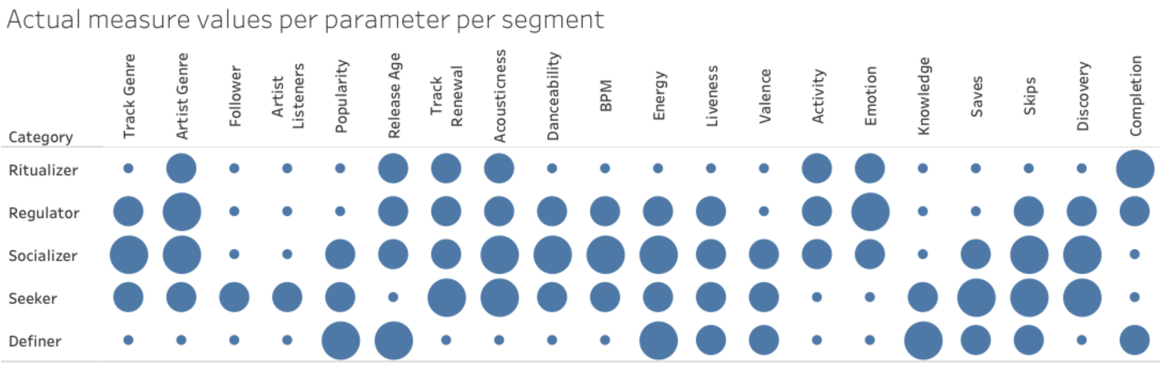

Fig. 3.16 Insight: Absolute Measures

The manner in which behavioral metrics are displayed in 3.16 makes it possible to observe a symmetry that allows for further categorizing the audience segments. This symmetry is in one case very apparent if one exclusively looks at the Knowledge and Activity metrics. Although the audience segments at the lower end of the spectrum display a high degree of musical pre-knowledge, the upper ones do not. In addition, the segments at the upper end of the spectrum display a high degree of activities accompanying the action of listening, and the lower ones do not. The same applies to emotional attributions. This symmetry can be escalated to the Completion and Discovery metrics among others. The completion rate can be seen as an intersection between interaction and relation indicators, as it is in accordance with activity-based and emotional attributions. When looking at the behavioral attributes of each listener segment, the following highlights are accentuated. The Seeker and Definer display an averagely high interaction potential combined with high degree of knowledge.

This stems from their approach to initiate a listening session with the main incentive of actively engaging or listening to music. On the basis of this primary activity, discovery, save and skip rates are higher in contrast to the other segments. This goes alongside higher pre-knowledge about the musical content being required to find, understand and enjoy the content. When comparing the Seeker to the Definer, we see a clear difference among the metrics Discovery and Skips. This is due to the difference in incentives when starting a listening session. The Seeker aspires to be exposed or expose himself to new content and thus to discover within a certain scope. Whereas the Definer defines more precisely what content he wants to discover beforehand which results in a partially limited discovery scope. Additionally, the overall amount of new, unknown or very different content is more limited among the Definer segment, which results in a medium degree of Skips.

The query calculating the degree of differentiation ultimately supplies an answer to the question of by what degree (between 1 and 3) can a parameter of one audience segment be differentiated from the same parameter of the remaining four segments. This is conducted on the basis of statistical validation metrics, derived in the statistical analysis section. Those are ultimately concatenated by means of an index that presents the power of differentiation per parameter and segment. This is calculated by means of the relative importance index (RII), which concatenates the absolute scores of all available metrics for each parameter and thereby induces a ranking of each segment:

$$RII = \sum \frac{W}{A \cdot N} \qquad RII = \sum_{i=1}^{3} \frac{W_1 + n_2 \cdot W_2 + n_2 \cdot W_3 + n_3}{3 \cdot N}$$

W = weights as per Likert's scale in a range from 1 to 3
1 = negligible impacts, 2 = moderate impact, 3 = major impact
A = Highest weight (here 3)
N = Total number of results in sample (all dimensions and underlying measurements)
n = Number of results per feature

Those indexed scores refer to high potential starting points and can assist much like a score card to gain an overview when many metrics weigh into the evaluation. Among the integrated metrics are the behavior of minimum, maximum, average, deviation and distribution and variability potential of a metric in the context of all five audience segments. Those are ultimately rated with the help of the degree of power of differentiation. The higher this degree, the earlier in the process the analysis accommodates those characteristics and the stricter the thresholds. Ranging from 1 to 3, a score index of 1 describes a low, 2 a medium, and 3 a high potential of distinction for the given parameter.

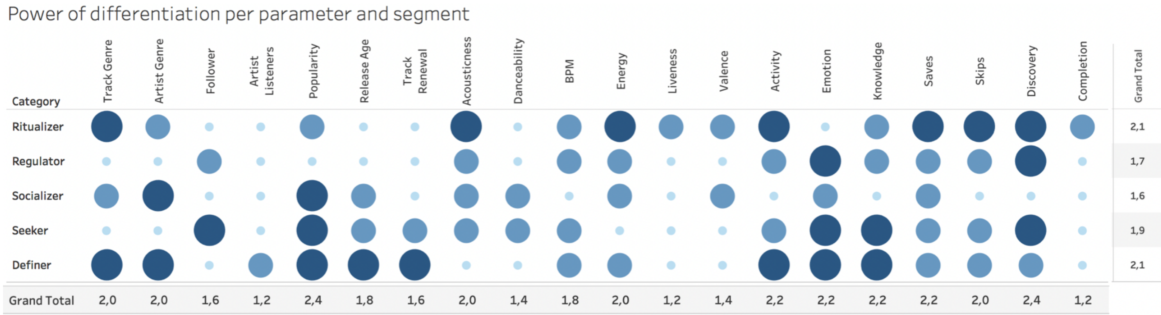

Fig. 3.17 Insight: Power of Differentiation

The audience segments that display on average the highest index of differentiation are the Definer and Ritualizer, both with an average index score of 2.1, which indicates a high level of distinguishing characteristics across all metric categories. Therefore, those segments are the most clearly definable and discernible from the others. Those are closely followed by the Seeker, with an average index score of 1.9, and Regulator with 1.7, rounded off by the Socializer with 1.6. Looking at it from a property-centric perspective, the framework discloses that Popularity, Activity, Knowledge, Emotion, Discovery, and Saves have the highest power of differentiation and thus are the factors with the strongest potential to differentiate between the listener segments.

3.3.2 Interpretation & Discussion

As per the hypothesis of this thesis, the central quest is to answer whether the establishment of a newly revised audience segmentation method, focusing on consumer- and context-centric data points, can facilitate a better understanding of the audiences' demands by analytical means. To achieve this, we aggregated data points to substantiate a newly derived listener typology, based on empirical research. The results of the statistical analysis on the previous pages show that the aggregated datapoints, including consumer-centric and contextually connotated parameters, allow for establishing distinguishable parameters for the five audience segments, allowing deeper insights into their habits and needs in regards to music streaming. The statistical analysis allowed for aggregation of the raw data that focuses on a multitude of findings. Those aspects are characterized by their high degree of discernibility or congruency from a standalone perspective, per parameter and single statistical validation metric. This proves that raw data are insufficient to derive actionable insights on the listener, even if sales metrics are removed. One reason for this is that despite providing valid calculations, the results are not put into relation. This changes with the aggregation of absolute measure values, as patterns are detectable and an overview can be derived. This approach ties back to the established hypothetical listener types and verifies its assumptions. When looking at the absolute measures (see Figure 3.16), two unexpected exceptions were outlined by the data, which demand amendments of the written definitions of the listener types. First, the highest average scores for an artist's popularity are found in playlists catering to the Definer segment. It has been stated that especially Socializers aim to listen to popular content. Second, valence was assumed to be a clear aspect of discernibility and thus higher in the Socializer segment than in others, but this ends up being just as high as in the Seeker and Definer segments. Both insights result in decreased differentiation potential for the Socializer segment. Ultimately, although those singular aspects of the hypothetical audience segments were contradicted by the analysis, the basis of the framework remains.

However, by opposing the results of absolute measure values and the power of differentiation, it becomes apparent that all metrics should be seen in a context of various statistical validation metrics for the observed parameters. This is required because when looking at absolute measurements in relation to other parameters or segments, on a numerically equalized level, their insights are at times proven wrong or insignificant, despite providing tangible insights. Further, the manifold parameters outline the importance of focusing on a limited set of characteristics that allow for differentiating those segments from another. This idea is implemented by the power of differentiation, which is derived by concatenating and indexing parameters per listener segment. This allows for the derivation of collinearities and synergies between different aspects depending on the contextual settings. It moreover results in so-called third degree indices and conveys the factors that are most distinct and therefore potent at discriminating between listener segments.

On the basis of this ultimate approach, power of differentiation discloses that the strongest discriminating parameters across segments are Popularity, Activity, Knowledge, Emotion, Discovery and Saves (see Figure 3.17). Five out of those six parameters are rooted in listener-centric metrics, which puts an emphasis on the established theory that it is important to include and stress those over product or sales data. With this approach, the choice of a consumer-centric approach, considered crucial by empirical findings in the fields of psychology and media science (see Chapter 2), can be endorsed. Although the analysis outlines that interaction and relation features display the highest potency overall, those relate to different parameters per segment. Thus, indices provide highly individual start points for content curation and analytics per segment. Furthermore, each segment has its unique total score. This shows that power of differentiation leads to knowing the most about the listener preferences of the Ritualizer and Definer segments. The power of differentiation can thereby be seen as an instrument to transferring the psychological and sociological knowledge on the consumers' needs and preferences into a measure that prevalent analytical tools miss. Thus, index scores provide a streamlined access point for understanding differences between listener types and allow for keeping the soul in the data by pairing human interpretations with quantitative measurements.

Consequently, the indexed power of differentiation allows for combining all of the previously derived insights. It can be concluded that using the suggested methods of deriving and processing the data can provide more tangible insights than traditional analytical methods. Another critical aspect that contributed to the choices of the parameters and ultimately leads to the unequivocalness of the index scores is the outline of the five listener segments. Commonly, approaches in industry-related music analytics tend to first arithmetically assess outliers before incorporating socio-psychological indicators. In contrast, the here-derived framework was established in an empirical manner based on the detection of deficiencies in existing research in sociology, psychology and media science. The process was turned upside down, conforming to the manner of research approaches typical for social science and humanities, which were thereafter substantiated and validated by data analytical insights.

The main conclusion that can be drawn is that there exist universal, stereotypical listener profiles for each segment. Those have the capability of facilitating the detection of preferences in music content selection and curation per audience group. The differentiation parameters are the most distinct among consumer-centric metrics. The overviews combined with the insights of the typology allow for improving the understanding of who a listener is and what he prefers or dislikes, as well as what his motivation is for listening. On the basis of those given analytical insights, the combination and cross-referencing of multiple layers of the metrics allowed an unambiguous differentiation and substantiated the five constructed audience types. Other metrics can thereafter be added for in-depth analysis, though the examined basic metrics establish the foundation that indicates which audience segments the given listeners should be attributed to or vice versa. Moreover, Figure 3.17 provides simple access points for comprehending differences between the listener types as summarized in Figure 3.18. This allows for understanding the differences between audience groups and determining problems or bypass such in application cases.

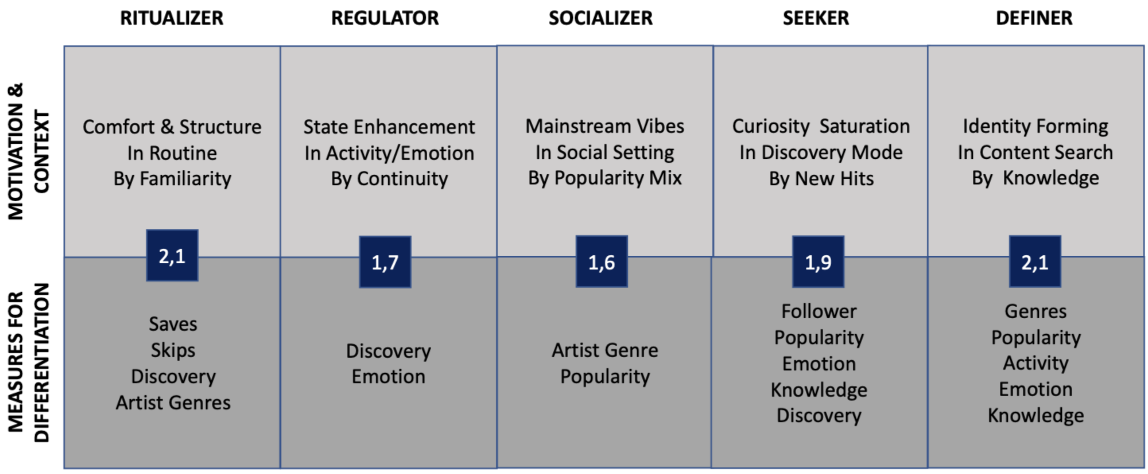

Fig. 3.18 Summary: Access Points for Audience Segments

Alternative explanations for the findings have been discovered in three areas. First, the insights to the absolute average measures as well as the power of differentiation can assist in deriving sound profiles for each segment. For example, a basic configuration of a sound profile may include a combination of specified sound features such as Acousticness, Danceability, BPM, Energy, Liveness and Valence which may be extended as desired. Applied to a smaller scope, this can especially assist in refining and shaping sound spheres. Second, the findings indicate that in the empirical as well as the analytical scope, demographics may not have as great an impact as in recent decades. This incorporates elements of expansion and increased de-centralized popularity of urban genres, driven by collaborations of artists as well as the lowering of accessibility issues due to an increased digital distribution in the music landscape. Third, the variety of the examined universal listener profiles offer a range that can adapt on the basis of the situational context and preferences of a listener. A consumer is not bound to one taste profile but instead can have multiple identities. The research findings illustrate that if a framework is desired, the main criteria of distinction need to be flexible enough to adapt to the given circumstances. This has been found with the listener's intention and the adhering context. For instance, a listener can consciously decide to listen to familiar music in a laid back manner as typical for the Ritualizer, however after a while display preferences for continuity while aspiring to enhance a state the listener finds himself in. Thus, oftentimes in a matter of seconds, scenarios such as a shift in mindset and behavior while listening to the same content or a switch to content curated for another audience segment, cause segment changes according to the circumstances.

In the following, some elements explain how this study is differentiated from previous research models. Systematic studies of music requires the interaction of several methods when deriving classifications and typologies. These are often based on term pyramids with genus proximum and differentia specifica. Those are not present in the here-derived typology, as the listener segments are aligned on one height and the flexibility of contextual states needs to be taken for granted. In addition, it must be mentioned that the here-developed music-sociological model is designed in the style of an ideal-typical model. Thus, no insights refer to regions, genres or social subcultures that would focus more on particularities. This circumvents the problem wherein sociological surveys and theories are frequently relativized by social change. Instead, context and motivation display a longer durability, because they are not tailored to a specific niche audience but the full listenership, where changes come into effect only after some time has passed because of the indolent system. Approaching this with a top-level, universal model results in a typology geared toward profiles free from any categorical boundaries such as genre or cultural taste hierarchies. Such a top-level model assists in putting a listener's individual profile into perspective.

Once all demographic, socio-graphic and music-related restrictions have been stripped, an individual's listening behavior can be observed in regards to his/her unique wants and needs, allowing for a broader music consumption than anticipated. Because this research aims to enhance typologies and not personalization, those criteria are assembled in five superordinate groups, which provide a stable basis for further specification into subcategories if desired. This is in contrast to prior typologies within the fields of music psychology and sociology aimed at parsing a listener's main activity. One example for this is the typology by J. Sloboda et al. (2012) in which six segments had been established. Those are travel, physical work, brain work, emotional work and attendance at live music performance events as an audience member. The here-derived results prove that the character of a playlist is not determined by an activity but rather the listener's intention and motivation while performing an activity, which is summarized with the listening context. Furthermore, Sloboda et al. identified four patterns for music usage behavior: distraction, energy, entertainment and meaning enhancement. Those attributes display the need to integrate the listener's intent. However, the motivation and context were not associated with singular audience types. Moreover, those insights were mainly been derived from qualitative and ethnographic research, where test persons talk or write about their behavior (Sloboda, 2012). This did not allow for the typology to establish a measurable, categorical and replicable framework, which the here-derived research permits on the basis of its mixed method approach. According to Clarke and Cook, "nominal and ordinal data, and their associated statistical analyses, have so far only rarely been seen in empirical research applied to music. Instead, most studies collect and analyze continuous data" (Clarke and Cook, 2004, p.202). However, the here-derived method of parsing consumption behavior makes it possible to reconstruct those otherwise un-observable correlations in a quantitative manner.

In regards to limitations, the following elements have to be observed critically. Even though the measurement insights for Skips can be revealing, the timing of measurement has to be treated with caution, in light of substantiated data capturing. Skips are only reported to the content owner after a track has been listened to for 30 seconds. Because this duration is in most cases already an essential proportion of a track's length, this parameter needs to be interpreted on the basis of the assumption of how likely someone is to still skip forward if they already listened to the first part of a track. This offers one statement on why the skip rate is lower for the Definer than for Seekers and Socializers, even though its listeners overall have a higher interaction rate and consciousness. Such a predisposition needs to be observed critically as much as evidence-based interpretations of findings in regards to relation indicators, which are activity-, emotion- and knowledge-based. This is due to a determination based on natural language processing, which handles titles and descriptions of the playlists. Those are meant to facilitate leading the listeners to suitable playlists. Thus, those metrics cannot be understood as factual measurements of emotional arousal, activities performed or musical knowledge of participants because the analysis does not track a listener's behavior outside of the streaming platform, nor inquire about it. However, it does imitate the search process a listener has to take on to get to such playlists and decisions he has to make before choosing the content. This approach is affirmed by Spotify's browse section, which allows users to quickly choose playlists on the basis of thematic pools, such as mood, genre, decade or activity. Descriptive data and behavioral metrics allow for the highest degree of approximation to the actual listening context and motivation. Trackable devices that gain insights into a listener's daily activity, emotional states and musical pre-knowledge are not feasible within the proposed outline of this thesis. On the one hand, it needs to be considered that the insights are generated without inquiring interviewees, which because of the surveying nature has a lower number of respondents but reflects assignable answers to specified statements. On the other hand, quantitative data, in addition to reflecting the technologically traceable content, enables a close approximate to actual states and a derivation in high quantity and is furthermore reproducible. In consequence, this limited provability of factual executions for relation indicators will pose a challenge for future research.

In addition to the listed limitations, the Regulator segment discloses another constraint that needs to be treated with caution. This segment receives the lowest score because of low importance rankings among parameters where all other genres are able to set themselves apart (see Figure 3.17). In regards to the 1.7 score of the Regulator segment, it needs to be noted that 17% of playlists contained in this segment are of algotorial nature. Those present partially curated and algorithmic tracks personalized to each listener's preferences. Thus, the discovery rate is elevated to a third degree feature, and there are higher skip rates. It can be expected that both would be lower under normal circumstances. Furthermore, insights into the dispersion of values within singular playlists are notable for this segment, as significantly higher constraints become visible from that perspective. Furthermore, it needs to be highlighted for this element that the overall wide coverage of degrees per parameter display only the options for variance per playlist. However, this segment discloses the importance of additionally observing the parameters on a per-track basis. This displays how narrow the levels are within a playlist, oftentimes covering only 1–3 segments, especially for Energy and BPM.

Recommendations that can help expand and reinforce the here-established research results concern an awareness of the framework's limitations. This could be solved by integrating real-time device tracking of activities or tests to further investigate music vocabulary to substantiate the evaluation of musical pre-knowledge with results from qualitative research. Moreover, music streaming analytics oftentimes include more audio properties such as Instrumentalness, Loudness and Speechiness. Those have been excluded since the here-derived statistical analysis does not differentiate between different musical genres. In this context, those three factors might have delivered distorted results, and moreover, the latter two are oftentimes subject to the audio mastering. Furthermore, the exclusion or conscious inclusion of playlists with an algotorial nature could lead to discrepancies from current results. Despite those recommendations, the here-derived framework and indices are capable to provide sound guidance for application cases, as outlined by the three case studies in the following section.

3.4 Application of Framework

The derived empirical framework was applied to multiple industry-related scenarios to prove its applicability for audience segmentation and to test the established indices. The derived results were conceptualized in a typology that allows to cater to industry-specific cases, such as label-, product- and brand-owned music content. For all those scenarios the same backbone has be re-purposed with minuscule amendments.

The metrics depicted in the previous sections allowed only for deriving encompassing insights on a superordinate level, without going into details of specific tracks, artists, brands or labels. But for the following scenarios, content ownership enables the provision of supplementary perspectives because it allows access to additional data pools as provided by the Warner Music sublabel Spinnin' Records. Thus, when applying the framework to industry scenarios, content ownership allows for complementing the so-far aggregated consumption-centric metrics with sales metrics, listener demographics such as age, gender and location, and social media insights as per availability.

Because any scenario would need to have an initial knowledge base as a starting point, the three exemplary scenarios each present one of three pre-defined scopes. Whichever element imposes the highest degree of restriction, audience-, context- or product- based information, was selected as the main element, called seed. The information about the seed is enriched by consecutive concatenations of those three categories as per known compliance needs. Once the content was filtered on the primary seed level, further additional restriction may be taken into account if known. This is essential because end- products are frequently shaped by multiple factors of the given three categories. Ultimately, the options for primary information are the seeds for the three correlating implementation scenarios.

	Seed	Exisiting component	Resulting component
1	**Product details**	Product	Audience segment
2	**Audience details**	Audience segment	Product
3	**Context details**	Product & Audience Segment	Enhanced Product

Table 3.3 Application Cases

The derived audience framework and indices enable the answering of different aspects for each of the three conditional states. The resulting components of those three implementation scenarios investigate the following questions. First, who is the product audience or is product being targeting to the best-suited audience? Second, what is the ideal product configurations for the audience? Third, how can the product curation be enhanced for a specified audience?

Those scenarios can be used not only for analytical purposes, see example 1, but also retroactively for creational processes, see example 2. At times, both goals, creational and analytical, are combined within one query when enough seed elements are given, see example 3. For all of those, the funneling process is facilitated by the derived framework and indices, because it provides clear indications for the starting point and guidelines for further refinement. At baseline, weights of the indices imply what parameters are the most reliable and should thus be focused on per audience segment. Moreover, framework and indices streamline the audience segmentation process significantly and furthermore incorporate so far not captured indicators of behavior and interaction on a context-specific level. If required, those scenarios additionally showcase what to focus on when incorporating other data streams.

The following steps are required to concatenate all necessary data streams and benefits of the provided framework. First, the seed needs to be defined, which can be one detail of the context, audience or product segment. This determines the primary filter for the first query. Second, further restricting factors, which can be one or multiple details of the context, audience, product segment are defined. Those specify the filter factors for the second and third query. Third, the factual population is defined by observing the total and percentage share of a subcategory's population within the target population. Fourth, the sample population is delimited by extracted data points with the same dispersion from the token dataset. If necessary, the selection of utilized user tokens needs to be amended to have the best possible representation of the total user base. Last, extending data sources can be consulted to refine the results. Those include market benchmarks and social media metrics as well as sales and stream numbers and region specific trends.

In regards to the seed and restrictions, the respective mandatory data-points can be superordinate factors of a desired query. The definition of a sample population can best be derived by questioning who exactly needs to be targeted by superordinate demographics, such as gender, age and location. If the primary query targets a large population, further partition can be provided by running a second query that looks into more behavioral and acoustic parameters such as music consumption source, genre, mood and activity. If multiple of the latter parameters are differentiating factors, a third query has to be added. For a scenario where a lot of knowledge of the audience prevails, all three queries would be subject to factors regarding audience metrics. In contrast, another scenario might present the context and a rough outline of the genre, so that one query regarding context and a secondary regarding acoustic features would be called upon. If too little knowledge is available, a query targeting the total volume of a streaming platform can be tapped. In regards to the sample population, the acceptable variance depends on the sample size. For all calculations, a confidence level of 99% and a margin of error of 5% has been placed. The sample size to provide statistical relevancy, is calculated in the following manner:

$$\text{Sample size} = \frac{\frac{z^2 \times p\,(1-p)}{e^2}}{1 + \left(\frac{z^2 \times p\,(1-p)}{e^2 N}\right)}$$

N = population size e = Margin of error (percentage in decimal form) z = z-score

Once a sub-group of a listener segment exists in a concentrated form, assembling large enough sample sizes becomes possible, which allows for analysis. After establishing this precondition, the examiner has to consider what percentage of all users opted in for the data collection and, if required, amend the selection of utilized user tokens. This is possible because the overall percentage of the user-based demographics is always visible in analytical tools of streaming as well as social platforms. For example, if both genders were to be considered for a query and the overall gender split is 80% male and 20% female but the tokens display an equal ratio, then a manual selection of the utilized tokens would be required. If the sample size comprises too few samples, the resulting data is not representative of the target population and because of this inaccuracy is unable to inform decisions. Depending on the case scenario, the sample size formula needs to be consulted for larger or more defined target groups. In the following, the analytical integration an application of the framework is contoured.

Case 1: Product Details Known

Case:	Gaming by Spinnin‘ Pixel
Goal:	Define audience segment
Question:	What audience segment do the listeners belong to?
Restricting factors:	Context – Gaming; Genre – EDM and related.

In the first example, the product details are given as the seed. The audience needs to first be understood to know to whom an audio product should be tailored. This process can be facilitated by assigning one of the derived five listener types. The exemplary Spotify playlist, called Gaming by Spinnin’ Pixel, is aimed at people who listen to music while playing video games. A problem arises when trying to allocate the listener type, because the motivation of a gamer who listens to music could be either of two types. The first type aims to focus on a reinforcement of the current state, as typical for the Regulator segment. This would further include a focus on low-level audio features because music is consulted as a background activity despite having a high level of cognitive alertness. The second type aims to enjoy games primarily as a social activity, as typical for the Socializer. Thereby, the constraints are less restricted and can vary between high- and low-level audio features as the attention shifts between background and foreground listening.

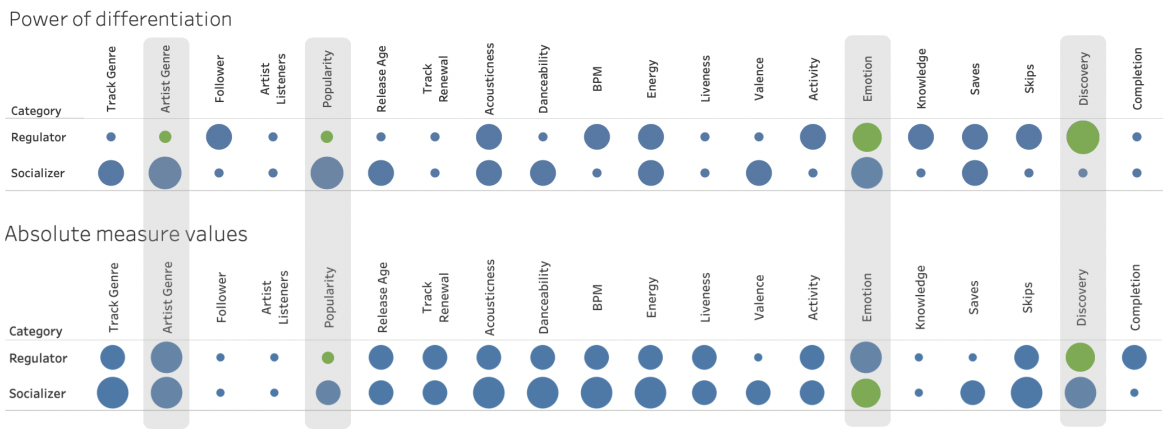

Fig. 3.19 Comparison: Actuals & POD - Regulator & Socializer

The index scores of the framework indicate which parameters have the highest potency of differentiation and thus need to be consulted 3.17. For this, the framework displaying the power of differentiation allows for identifying the parameters that need to show compliance.

These are the ones with the highest indices. Thus, all third degree indices of the Ritualizer or Regulator segments are highlighted in grey and are benchmarked against the exemplary playlist (see Figure 3.19). The final segment of the product can be identified depending on which criteria of differentiation are best fulfilled, highlighted in green.

To obtain the marked results, the questions are as follows: First, which are the third degree indices of the listener segments in question? Second, how important are the highlighted indices for the distinctiveness of the product? Third, which of the highlighted absolute measures is closer to the one of the product? Fourth, which segment scored the most points across all selected parameters, considering indices and absolute measures? It has to be noted that an analysis in an artist or label context demands all parameter be subject to the respective context. Thus, a high or low average of a parameter needs to be observed in relation to all of Spinnin' Records' playlists. Thus, the three sizes in the absolute measure framework indicate a performance above, in or below Spinnin' Records' average. In the case of Pixel, once the indices of the Regulator and Socializer segment are observed, the following third-degree parameters are established: Artist Genre, Popularity, Emotion and Discovery. After assessing the parameters of the playlist, it is shown that all indices correlate with the importance weights represented by the Regulator. Thereafter, the absolute measures are consulted for instant benchmarking of average figures. Thereby, Popularity and Discovery are indicative of the Regulator, Emotion indicates the Socializer and Artist Genre is a tie.

One element that naturally impedes this decision-making process is that the Regulator and the Socializer are adjacent segments within the typology. Therefore, many of the parameter values display only small differences. This emphasizes why the power of differentiation needs to come into play. If one looks only at the absolute measures without consulting the parameters with high indices, it is observed that almost half of the parameters indicate prototypical Socializer tendencies. However, when focusing on the most differential parameters, this distortion can be eliminated. Once the averages are compared, the standard deviation, outliers and range associated with the high indexed parameters of one or both segments allow for further comparison. In this second phase, the Regulator segment is yet again more supported than the Socializer segment, as highlighted in green in the initial row. This becomes visible by looking at the high index parameters of the Regulator segment, which correspond with the current product configurations and consumption patterns within this particular playlist.

Moreover, when consulting the listener motivation, further aspects endorse the assignment to the Regulator segment. First, the clearly defined scope of BPM with all tracks invariably on 120 BPM allows for keeping up the heart rate and reinforcing the current state. Second, the Energy levels are exclusively between 75 and 90, not allowing for large interruptions of the ongoing flow. This differs from Spinnin' Records' average measures for Energy levels, which generally vary between 50 and 100 with a bell-shaped distribution. This leads to the end results that despite Socializer tendencies at first sight, the indices with their power of differentiation allow for additional insights disclosing that 75% of the assessed parameters indicated an affiliation with the Regulator over the Socializer segment. Further curation with the preferences of this segment in mind can be undertaken to increase this share.

Case 2: Audience Details Known

Case:	Spinnin‘ Records Fit
Goal:	Create playlist for specific audience
Question:	What are the ideal product configurations for this audience?
Restricting factors:	Segment – Regulator; Context – Running; Demographics – Spinnin‘ Records.

In the second example, the audience is given as a seed. To know how the product configurations should be defined to cater to a specified audience, first an external context scenario needs to be chosen. In this scenario, the overarching goal is to set up a workout playlist and tailor it to a pre-defined audience. For this, the two preconditions are that it has to cater to listeners of Spinnin’ Records and that, among those, listeners of the Regulator type are targeted. For this purpose, the following case study showcases how the derived segmentation of the Regulators segment can assist in audio feature profiling, on its own as well as in combination with demographic and acoustic pattern analysis.

Analyzing the acoustic parameters takes place from more general to more granular levels. The creational process of any new playlist within this segment could look like the following. First, the scope within the Regulator segment can be limited to a range of activities including workout, sports, running and outdoor sport playlists. Second, all playlists that fall into the given scope are screened for seven acoustic parameters over months on a weekly basis. Those include the acoustic features Energy, Valence, Acousticness, Instrumental, Speechiness and Liveness.

This investigation with a focus on acoustic parameters allows for instantly summarizing that all playlists within the concerned activity and context scope introduce Energy as the highest level throughout (avg. 80), followed by a high level of Danceability (avg. 60) as well as an omnipresent level of Valence (avg. 40). Figure 3.20 shows five examples of playlists within that scope where the values of their acoustic metrics lie within those averages, with all included tracks taken into consideration. The main concept of all five playlists is inherently the same: Cater to people who take up workout activities, which is a representative scenario for listeners of the Regulator segment. In addition to overarching similarities in context, motivation and some acoustic parameters, it has to be kept in mind that the purpose of these five playlists is fundamentally different. It ranges from artist exposure to brand marketing and event promotion for at times more or less defined audience groups.

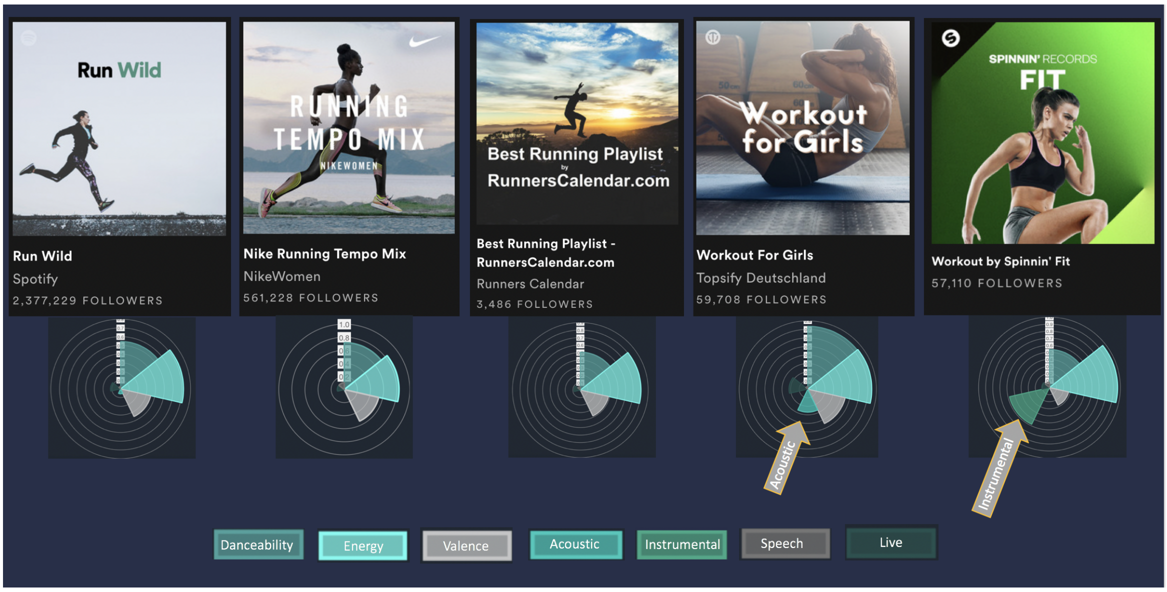

Fig. 3.20 Outline: Acoustic Features - Workout Playlists

The varying objectives have an effect on the audio profile because of the impact of accompanying secondary restrictions or preferences. Secondary queries are necessary to shed light on subliminal factors that help to further determine the product on the basis of restricting factors of the targeted audience. Those requirements are most often determinable by the specific demographics of the target group, though also connected to psychographics or specific acoustic determinants. Those secondary queries in the specific use cases displayed are as follows: Run Wild by Spotify, no secondary queries explicitly necessary; Nike Running Tempo Mix by NikeWomen, 2nd women and 3rd Nike Runners; Best Running Playlist by Runners Calendar, hits of sales category; Workout for Girls by Topsify Deutschland, 2nd females in Germany; Fit by Spinnin' Records, 2nd Dance and EDM genre and 3rd brand-centric Spinnin' Records.

Those secondary queries specify audio features by which a particular playlist sets itself apart from others within the same context segment, based on subquery limitations. This allows, for example, defining that females between 18 and 35 years old prefer to listen to tracks with a higher Acousticness factor, which is among other playlists also present in the audio features of Workout for Girls. However, this outlier can be excluded in this case scenario by directly sub-querying for Dance Music. It then becomes visible that Dance Music fans display another case of exemption that, for instance, also present in Spinnin' Records' Fit playlist. Hereby, the instrumental factor is very dominant. Additionally, the relationship between Energy and Danceability is, with a ratio of 1:2, greater than the group average (see Figure 3.20).

Consequently, the first query provides an overview of an acoustic profile based on the selected audience segment, and the subqueries highlight in detail on which acoustic factors the specified audience sets itself apart from the generalized profile of the segment. When contrasting all playlists from the highest level of selection to highly specified and restricted queries, different specifications become visible among audiences which underlie the same constraints. Thus, at times audiences display diverging preferences while being subject to the same audience segment, which for the here-derived example is the Regulator segment. In this case, the motivation and context is shaped by the enhancement of an activity that is further characterized by continuity in musical parameters (see Figure 3.18). Moreover, tertiary queries offer an option if further refinement is wanted and additional elements of the seed are known.

For instance, an ideal-typical process to set up a playlist such as Fit, with the main limitations being the genre scope, age group and gender, could look like the following. While neglecting the prior existing playlists, the procedure would take place in a three-step model. After looking into the entirety of workout playlists, subquery factors allow limitation of the scope by querying first for Dance and Electronic Dance Music genres and second for demographics limited to 18–35-year-old males. Accordingly, the secondary queries allow for more granular targeting, which assists in further refining the tracklist.

Those steps do not require knowledge on the audience segment. However, curational steps on the most granular level do require insights into usage patterns and the intrinsic listening context and motivation, because those concern the selection and organization of tracks. The goal would be to not forget about the listener's main motivation for listening despite so far having focused on genre and demographic restrictions. With the knowledge on the listener segment, the curational process can be facilitated by consulting the metrics with the highest index of the given segment, Discovery and Emotion (see Figure 3.17). In the case of the Regulator segment, with the help of the indexed framework, it can be immediately detected that discovery potential and emotional attributes need to be curated with care because they display highly potent third-degree indices for this segment. In contrast, Popularity and Valence, among others, can be neglected, as they display low potency. Further, the absolute measure (see Figure 3.16) can assist in benchmarking whether averages align and if applicable to amend the parameters according to the standards of the delimited sample.

The consideration of the listener segment allows for bringing the act of listening and the surrounding experience back to center stage. This emphasizes the differentiation between the obvious external context that is equivalent to activities, states or daytime as given with the seed in this case study. On the other hand, internal context is comparable to intrinsic motivation and a listener's lean-forward or lean-back character throughout the act of music listening. Although the external motivation can be singled out in queries and facilitate the derivation of acoustic profiles, the internal context outlined in chapter 2 enables the curation of the content. Thus, before a playlist is created on the basis of the knowledge of an audience group, high-level queries should be run to understand consumption preferences within a contextual setting; second, subqueries can be added and, third, the consumption preferences of the listener type are interwoven during the curational process. Thereby, creative decisions can be taken that may lead to the choice to assimilate existing playlist profiles or to newly develop a unique profile that sets itself apart.

Case 3: Context Details Known

Case:	Spinnin‘ Records Brand New
Goal:	Enhance playlist curation
Question:	How can the content for the assigned segment and product be better curated?
Restricting factors:	Segment – Seeker; Context – new releases; Demographics – existing audience; Acoustics – per label.

In the third example, the full context encompassing audience and product are given. Existing playlists can be revised or readjusted on the basis of the current performance of playlisted tracks compared to the ideal-typical parameters of the given audience segment to know how the content curation could be refined to better cater to a specified audience.

Plenty of existing data analytical procedures aim to evaluate performance of music content based on streaming performance. However, although those analytical processes at times include the monitoring and assessment of behavioral metrics, re-evaluating why values are considered to be a positive or a negative trait has oftentimes been neglected. In general, high Skip rates are considered a negative trait, and high Save and Discovery rates are considered a positive trait. Beyond those numbers, users exist with their individual motivation and intention for listening that shifts the perspective from which the interaction parameters are being assessed. For example, for a Regulator to have a low Save rate does not indicate that the product performs badly. This is because the Skips are plainly low for this user type as defined in the table of absolute measures. However, on the basis of the index that encompasses all metric descriptors, the Regulator is able to set himself apart with his third-degree index on Saves. This power of differentiation indicates that one should focus on this parameter. Observing the behavioral parameters in Figure 3.17 demonstrates that especially five aspects should not be overlooked during evaluations. Those include Saves of the Ritualizer segment, Skips of the Ritualizer segment and Discovery of the Ritualizer, Regulator and Seeker segment. Those are respectively either higher, lower or more or less deviating than a prototypical playlist of the other segments that makes them more potent in the case of differentiation. This branches further out into the product and context parameters, which make it once more difficult to assess whether a high degree Popularity or Activity is to be considered as something positive. To answer such questions the table with absolute measures can be consulted. Last but not least, the acoustic parameters can be taken into account, although they are only valuable on the most granular level and would need to be benchmarked against similar content as described in the previously discussed second-case scenario.

Accordingly, the steps to improve the assessment of those indicators within existing products underlying a predefined audience segment could come into place as follows. First, the motivation and context are identified, which need to be satisfied according to the user type. Second, the third-degree indices are identified that outline the parameters which need to be fulfilled. Third, the list of absolute measures that indicate whether the norm is a high or low average are consulted to re-assess the evaluation. Those steps build the foundation to enhance currently existing analytical processes by adding the knowledge of the power of differentiation. This ultimately allows for assessment of a playlist based on knowledge while considering the intention of a listener. As in the prior examples, all retrieved scores need to be put into relation of the playlist-specific environment, which is in that case all Spinnin' Records playlists and its users.

For instance, the playlist Spinnin' Records Brand New caters to the newest releases to listeners of this music label. It is aimed at Seekers whose main motivation is the saturation of curiosity while being in a context in which they are driven by a discovery mode and consuming new hits (see Figure 3.18). This segment outlines by means of the third-degree indices that the parameters Follower, Popularity, Emotion, Knowledge and Discovery have to be taken into account. From there, one can observe the absolute measure of the parameters chosen by the indices of the Seeker segment. Those are respectively compared to the values of the current playlist and help to assess assumptions.

At all times, the title, description, playlist cover and its contextual implications and descriptions are the only two aspects that a potential listener encounters before deciding to listen to a playlist. Thus, it needs to be communicated clearly what context and motivation a playlist aims to satisfy and to which restrictions it is subject. Matching expectations must be raised to outline how one of the workout playlists sets itself apart from the others. This alignment of expectations can ultimately increase listener satisfaction and audience retention. This fulfillment of the content's purpose, such as branding or artist exposure, can only be reached when the derived preferences on consumption and curation of a segment have been considered.

Parameter	Segment: Seeker	Product: Brand New	Segment- Product Conformance
Follower	Medium	Medium	High
Popularity	Medium	Low	Medium
Emotion	Low	Low	High
Knowledge	Medium	Medium	High
Discovery	High	Low	Low

Table 3.4 Comparison: Actuals per Segment & Product

The comparison demonstrates which of the third-degree parameters show conformance on its absolute numbers. Although the parameters Follower, Emotion and Knowledge display full accordance, Popularity is one and Discovery two degrees lower than the typology suggests. This agrees with Popularity scores that are lower than the average of Spinnin' Records playlists. However, someone who is knowledgeable on this listener segment knows that this can encompass new track playlists that have untapped popularity scores as well as Hit playlists that cause the overall average to rise. Thus, considering that we are in this case working with a playlist aiming to present new releases, the medium conformance of this parameter can be waved through. The lower-than-average Discovery rates set themselves significantly apart from the segment's suggestion. One might be that this is due to the interaction potential of the label- or genre-specific audience. However, this aspect has already been taken into account by benchmarking the existing playlist against the remaining releases of the same content owner. Further, a low Discovery rate indicates that people have already listened to those tracks elsewhere before hearing it within this playlist context. A low Discovery value is typically associated with Ritualizers or Definers, but the motivation and context linked to those segments are not congruent with the incentive of someone who desires a playlist that contains mostly unknown tracks. Therefore, biasing messages need to be circumvented, whereby the title and targeting suggest satisfying one listener type but the playlisted content suggests other. In the case of Brand New, an increment of the Discovery value can be achieved by renewing the tracklists more often and possibly reducing the number of tracks to guarantee that exclusively new tracks are included. Of course more parameters than the absolute average can be taken into account thereafter. However, for a first comparison to assess conformance to assess the results correctly, this step is sufficient. Furthermore, this scenario could incorporate the approaches of cases 1 and 2.

So far, many misjudgments in regards to parameters have been made because of generalized observations. Now, the separation into audience types allows an analyst or researcher to more appropriately evaluate the listener's habits or preferences and consequentially better satisfy them. Ultimately, the existence of a comprehensive seed allows for evaluation of the current performers and steering the navigation by increasing or reducing parameters that display non-conformance.

As the outlined exemplary scenarios show, the main benefits derived from the usage of those newly derived elements are the facilitation of profiling audiences despite not knowing the full product scope in advance, but only certain seeds. This is enabled by a pre-segmentation of the full audience pool. Further, the approaches are highly customizable with a modular system. Additional use-cases for which the three above-mentioned examples would come into play and benefit from consulting the listener types as well as the accompanying indices are in the fields of brand partnerships, music licensing, competitive analysis and artist profiling. This can enable, for instance, a brand partnership manager to attain a limited selection of artists and respective tracks tailored to the audiences' desires and campaign needs. Likewise, this can be applied to multiple artists or music formats in a comparable manner when deriving a competitive analysis. While industry cases display diverse applicability on feasibility of the framework, they can also pass knowledge back into the source material of scientific methodologies. Learnings of current user behavior and market demands can be incorporated into the manner in which research questions are formulated and examined.

Chapter 4

Value of Combined Scientific Insights

This chapter aims to present cases which outline why a multidisciplinary systematic musicology and industry analytics profit from the concatenation of analytical with humanistic approaches. First and foremost, this need is articulated by the label attributed to the act of listening. While on-demand listening is still practiced by audiences such as the Definer, nowadays, the term music streaming is slightly inadequate. This becomes obvious when asking why listeners play music, which leads back to the central aspects of listening context and motivation. Instead of music streaming, this act can rather be referred to as a digital music experience. This resembles multifaceted happenings, characterized by multisensory and multimedia properties. Thus, visual and auditory senses may be mobilized while music from streaming portals is embedded in video commercials or spacial designs. Thereby, involvement of music can range from conditioning or accompanying to being a background element as per required context. Accordingly, the term itself indicates the breadth of factors that take stakes in the creation and interpretation of digital music experiences, going way beyond the scope of traditional music streaming.

Moreover, music used to be mainly consumed in music-first contexts, such as concert-going or vinyl-listening, where the listener's main focus was directed toward the act of listening. In contrast, facilitated by the digitalization of this sector, there is a tendency for music to be presented parallel to appended non-music-related contexts or even subordinate to those. The first-mentioned occurs as motivations and incentives get broader, which leads to the addition of real-life contextual relations, such as activities. The last-mentioned occurs because of the increasing inclusion of music into products and services outside its main scope, which aim to profit from its impact on listeners. Such use-cases of music in scenarios that reach beyond its traditional fields of application resulted in an increasing amalgamation of the industry, science and art sectors. As those disciplines strive to include music into their strategy plans, a demand to understand listener consumption, behavior and perception of music consolidates. However, such an analysis is only possible if standardized comparable metrics are available, as suggested by the here-derived framework and listener typology. The premise of music streaming as a multi-sensory music experience influenced by multi-disciplinary factors makes the concept of combined scientific approaches indispensable. This postulates that the impact of many factors can best be assessed by incorporating some humanistic approaches into algorithmic analytics just as much as the other way around. Thus, both disciplines will allow for enriching their insights and consequently their explanatory power. This can ultimately increase their contributions to sovereign processes because of an increased understanding of its effects.

In interdisciplinary approaches, the enhancements in understanding behavior based on measurable outputs also have a positive reverse impact, whereby algebraic calculations absorb humanistic interpretations. This surfaces as components of human judgment and interpretation are kept in analytical investigations. This is mainly facilitated by taking three types of metrics into consideration: first, psychological and behavioral descriptors; second, hard qualitative metrics in the form of the absolute values; and third, soft quantitative metrics with the power of differentiation, which is based on human assessment and comparisons of segments. Thus, without neglecting the humanistic aspects, the here-conceptualized technique embeds data power in an industry that is currently lacking in timely, granular and transparent data reporting.

In addition to the mentioned concatenation of humanistic and industry-specific approaches, the choice of parameters needs to be considered carefully to deliver the desired insights. As markets and sectors amalgamate, questions in regard to audiences have to be formulated more precisely to understand what resource needs to be tapped. In this case behavioral, interaction and direct evaluation metrics are crucial factors to open the way to a more user-centric understanding. New technology enables us to track and understand the behavior of people in real-time without interfering during the data collection process. On the one hand, the underlying technology can enable researchers to ask more precise questions. On the other hand, hybrid data serves to keep a human-like spirit in the data. Thus, instead of asking people how they think they feel, quantitative data can add a second dimension that taps into what and how they actually react and act if one aims for a full picture to fine-tune products or services. This is implemented by allowing designers to switch between both research methods during rapid prototyping. Seemann foresees an opportunity for hybrid data to point to the future of a *smart insight generation* (Seemann, 2012). However, many innovations in the music industry are presently led by technology rather than user needs or individual preferences. This bares the risk that products and services are created that consumers do not want to engage with, especially user interface designs and supplementary content elements such as audio-visual experiences that are highly contextual, which calls for a human-centered design process.

These new measurement techniques have been derived as a consequence of the increasing context-driven paradigm in music streaming. However, in application cases, the approaches need to be differentiated between very granular and segmented. The first is present in context-sensitive algorithms that analyze a listener's behavior to an extent where his activities, time of day and emotions are being understood, which enables it to recommend music that matches the listener's context. This is performed on a non-conceptualized and very granular level, aiming for personalization (Prey, 2017). On the other hand, it is essential to see overarching connections between user segments by means of an audience partition to allow for targeted recommendations that are not limited by one person's interaction history but rather display high-level categorizations. Because of this generalization, the segmentation approach is consequently more applicable for audience assessments. As segmentations induce a shift to contextual parameters and a repression of hard demographics such as age, gender, social class or location occurred in the process of creating typologies.

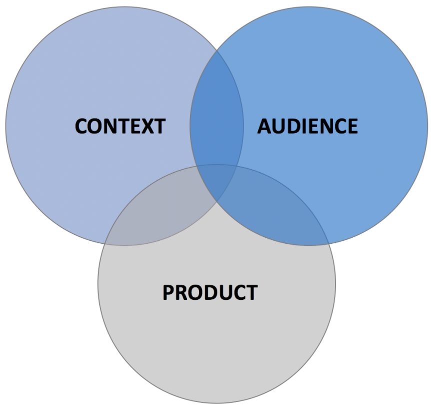

Fig. 4.1 Context Spheres

In the same manner, the selection of music increasingly detached from demographic preconceptions, after the digitalization of this sector. This asks for new concepts of audience categorization. Though, typologies drawn upon contextual descriptors demand more metrics that can be put into relation with subjective affects and outer influences. This is provided by observing a triad of insights, including product, audience and context spheres, where none of the three dimensions is superior to the others (see Figure 4.1). Such a concept allows for more fluid and flexible borders between listener types while allowing for them to break free and reorientate, thanks to the knowledge on user psychographics. On an individual level, demographics can still be consulted, for example, if a listener wishes to be notified about local upcoming artists depending on his location. But instead of imposing rigid structures based on demographics, those surface in the derived typology only as secondary factors. However, on the top level, the categorizations are regulated by flexible real-life scenarios as a listener may encounter them.

This approach is made possible by the amalgamation of scientific approaches. For one thing, disciplines belonging to the systematic musicology are at times jointly examined but only singular studies are evaluated with the support of extraneous disciplines that have capabilities to provide analytics based on user applications. While musicological efforts are able to derive the audience (who) and context (why), the configurations of the end-product (what) stays undisclosed. Application analytics can assist in disclosing this third factor. Furthermore, it discloses that audience and context do not necessarily have to be predetermined nor aligned. This becomes visible as the emphasis migrates from demographics to contextual descriptors. However, combined scientific insights allow musicological studies to continue examining music streaming, as they suggest a merger of those three areas by means of a listener typology incorporating all three elements. The segments of the typology establishes new research subjects based on the analytical partition in Chapter 3. This new subject allows musicological research to expand existing typologies by considering multiple factors at once. Apart from that industry analytics profit from taking the mindset off of micro- towards macro-structures. This requires analysts to forgo parameters which are ready at hand such as sales metrics and demographics for more seemingly intangible humanistic factors. This allows streaming analytics to gain insights into listening contexts that were previously immeasurable from a purely analytical standpoint.

These insights created by cross-disciplinary approaches are crucial to the success of music streaming platforms, as they rely on understanding customers and how they operate as consumers as well as people in their individual environments. This will increase in the coming years, since music streaming and brands are expected to form an even stronger alliance and evolve together. For instance, listeners have already displayed the desire to connect with their favorite influencers and brands. By treating music streaming platforms as a type of social network, brands have the opportunity to connect with their consumers on a deeper level and therefore inspire stronger loyalty among fans. For this the requirements for an audio product of the audience segment as well as of the brand need to be defined, before more granular content curation can take place. Two examples for the incorporation of personalized playlists into brand strategies are presented by Carnival Cruise Line and *Stranger Things*. Carnival Cruise Line has a Spotify playlist consisting of upbeat summer tunes that enable their customers to immerse themselves in the cruise experience. For *Stranger Things 2*, Spotify collaborated with Netflix and created playlists for the 13 characters based on its audience's listening history. The same applies to the integration of music into video games, retail experiences, virtual concerts, restaurant ambiences or advertising campaigns. In this way, music has the capability to extend the reach of brands to a level currently not exploitable with other social media channels (Treseder, 2018). Thereby, a varied array of disciplines need to be integrated to assess what music is best suited for those contexts and the respective audiences. Marketing segmentation can assist in understanding user demands from a market perspective. However, analytics enhanced by behvioral and psychological insights ultimately allow to understand users on a context-based level in so-far unrelated industry sectors. With the help of those insights, bespoke services, products and content can be developed or refined to satisfy corresponding user demands. In addition to customer satisfaction, the involvement of music can lead to a competitive advantage in a branded environment. Those benefits can be reaped by industry players such as labels, artist managers, music tech companies, tour promoters, radio promoters, brands and advertisers.

Consequentially, markets as well as sectors have to be open to interacting with adjacent disciplines. This spans across the full distribution chain and encompasses analytics. Thus, methodologies and parameters of different methods should be combined as well as knowledge insights. The necessity for this arose because music became an integral part of a vast variety of non-music-related products and services. This emerging cross-disciplinary relevancy can be captured in a streamlined manner by mediating the derived audience framework and indices. In doing so, multiple scientific methods as well as industry approaches are concatenated and can be mutually beneficial to either discipline. This newly derived typology allows all players to derive basic information about their audiences that can be further refined as per given demographics shown in Section 3.3.2. The consideration of all those factors allow for more tangible and reproducible insights. Thus, interdisciplinary, mixed-method typologies bring context, audience and product closer together, which can create added-value for as-yet unaffiliated sectors. Thereby, all three spheres should be considered as equally contributing parts to the end product and pivotal to understanding respective audiences.

Chapter 5

Conclusion

This research has investigated whether the establishment of a revised audience segmentation method, focusing on consumer and context data points, can enhance understanding of audience preferences by analytical means. Data points were aggregated to substantiate a newly derived listener typology, based on empirical research.

The key results indicate that due to significant changes of music consumption, listening behavior, and perception, a revision of the methodology is indispensable for assessing listeners in a contemporary and pertinent manner. Thus, novel listening preferences likewise demand novel assessments of the listening experience. A revised approach to audience comprehension is enabled by placing the listener at the center of every analysis before diving into quantitative metrics. Accordingly, user-centric metrics should be combined with sales-centric metrics to understand audience preferences. Further, the results indicate that the listening preferences are not primarily determined by demographics or activities, but rather the listener's intention and motivation while performing an activity, which is summarized with the listening context.

The renewed framework is organized around functional niches, catering musical as well as non-musical scenarios, and enhances existing models by adding consumer-centric and contextually connotated parameters. Those insights help substantiate the five listener types of Ritualizer, Regulator, Socializer, Seeker, and Definer, via an analysis of acoustic, interaction, product, and relation features. The manifold parameters outline the importance of focusing on a limited set of characteristics that allow for differentiated listener segments, as facilitated by third-degree indices. Across listener segments, interaction and relation parameters carry the highest potency for distinction. The power of differentiation can thereby be seen as an instrument which allows to translate psychological and behavioral knowledge of consumer needs and preferences into a measure missed by prevalent analytical tools. Thus, index scores provide a streamlined access point for understanding differences between listener types and allow us to keep the soul in the data by pairing scientific interpretations with quantitative measurements. This allows to reverse the approach typical for existing music industry analytics that involve arithmetical assessments of outliers that are later observed in-depth. Instead, the framework conforms to socio-psychological research approaches and is later substantiated and validated by data analytical insights.

The established methodology provides means to better analyze music data to derive listener-centric insights, as correlation and indexation allow for a further substantiation and a clear delimitation of the user types. These methodological advancements contribute to filling the research gap by delivering a typology applicable to varying audience pools and making it possible to represent otherwise imperceptible, interdisciplinary correlations. This enables the establishment of distinguishable audience profiles with the developed typology. Thus, as per the hypothesis, the revised audience typology can facilitate a better understanding of audience preferences and ultimately contribute to the satisfaction of those.

In general terms, the derived typology summons the shift of two fundamental paradigms in music streaming analytics: first, the establishment of context and motivation as key constituents of an audience typology, and second, the supplementing of sales- with user-centered parameters. In this manner, the study offers mainly industry-related solutions, which are tested by means of a metadata analysis and illustrated by exemplary queries, based on consumption data. Accordingly, the research question is carried out in a hypothetical style, which enables the testing of possible query mechanisms of digital streaming services in an industry-centric way. As listening context and listening motivation play a significant role in all major changes in consumption behavior, they were established as base categories for the classification. Alike the search categories for playlists in Spotify, the thematic areas were divided up using natural language processing. A five-fold division allowed the assignment of each playlist to a descriptive context of motivation. This derived typology framework can be considered as a hypothetical test object, before its parameters are refined by analysis and result in the final framework (see Chapter 3.3.2). Finally, three case studies round off these chapters and show how highlighting individual query categories can point out the listening behavior of users more clearly. It has to be noted that while the empirical approach collects data from individual listeners, the results only reflect a summary across groups of listeners. Furthermore, the research is carried out in the style of a top-level model, whereby categorical restrictions are removed in order to override any restrictions and to apply them as widely as possible. In addition, the final design represents a clearly structured, ideal-typical framework, yet there are transitional forms between different types of listeners that are not listed here.

The chosen hypothesis has a high potential to tie in with general scientific theoretical discourses and, with its problem definition, refers to a fundamental research question that goes beyond boundaries of disciplines. Music psychology and music sociology provide a starting point for this concern, as these aspects are mostly overlooked amidst marketing-oriented handling of musical content. Studies on music behavior that combine scientific and applied analysis are limited, but music preferences and listener typologies can assist in making this connection.

Beyond the outlined approach, increasing the linkage to existing theoretical models of psychology could further allow to strengthen the music-psychological foundation of this hypothesis. In this manner, the here displayed analysis of secondary and primary data, could be supplemented by other methods that facilitate personality classifications in music. For example, the Big Five Test evaluates the weak or strong manifestation of openness, conscientiousness, extroversion, agreeableness and neuroticism, which are likewise essential factors for the specifications of the above derived five listener types. This OCEAN model has contributed to psychological, neurological studies such as *Ein anderer Ton. Das Hofer Modell* (Pöppel and Welker 2009). Thereby, procedures for determination of personality structures include not only the NEO Five Factors Inventory Personality Test (NEO-FFI) but also other standardized tests from the Vienna Test System for testing memory, attention, intelligence and emotional perception. Such extending testing methods allow to assess mental, emotional and social competences and may include: Grazer-Assertivitäts-Test (GAT), Sustained attention tests (DAUF), Raven's Progressive Matrices (EKMAN 60 Faces test), Prosody-Emotion-Test (PET), Evaluation of facial expressions (DTGS), Five-factor nonverbal personality questionnaire (FF-NPQ), among others. Further methods that tie into the proposed hypothesis are investigations on the physical level that measure the arousal generated by the affective involvement verified with the electrical skin resistance. On a neuroscientific level, neuronal correlates of music perception can be verified by mapping the electrical activity of nerve cells (EKP, ERP) (Schröger, 2005; Allesch, 1981). Moreover, the cortical encoding of desired and unwanted music can be investigated with the method of electroencephalography (EEG) based on left and right hemispheric processing of perceptions (Altenmüller et al., 2002).

Such tests could contribute to further ground and verify the derived insights on consumer segments. However, it has to be noted that their execution does not fully align with the established methodological idea, which implies that it is elementary to emphasize on data volume, consistency and a non-disturbing data acquisition process, to preserve a context close to reality. Nevertheless, the mentioned tests could provide scientific contributions to singular measurement factors of the listening preferences and allow to differentiate singular cases on an even more granular level. Beyond this, methods such as real-time device tracking and surveys based on mobile notifications comply with the methodological ideal and could tackle the limited provability for factual executions of relation indicators, by verifying activities, musical pre-knowledge and moods. Lastly, the application of preference indices can be tested in science and industry as long as data protection regulations allow the access of open source tokens. Based on such data, future research could implement the addressed shift from sales- to user-centric approaches through the application of the typology framework in analytical and creational processes.

The investigation describes and visualizes the distribution of listener preferences within the expanding market of music streaming. In doing so, it delivers a first, testable model for future industrial data queries in the field of streaming services, which discloses a data-rich field at the intersection of associated disciplines. However, the influence of digitalization on music listening should not be seen as an end in itself. Instead, by parsing preferences, data can be harnessed and allow to identify valuable insights into audience segments.

References

[1] Adorno, T. W. (1975). Typen musikalischen Verhaltens. In T. W. Adorno (Ed.), Einleitung in die Musiksoziologie (pp. 14-34). Frankfurt: Suhrkamp.

[2] Allesch, C. G. (1981). Untersuchungen zum Einfluß von Musik auf Puls- und Atmungsfrequenz. Zeitschrift für klinische Psychologie und Psychotherapie, 29(4), 353-382.

[3] Altenmüller, E. et al. (2002). Hits to the left, flops to the right. Different emotions during listening to music are reflected in cortical lateralization patterns. Neuropsychologia, 40 (13), 2242-2256.

[4] Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. Personality and Social Psychology Review, 10(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2.

[5] Behne, K. E. (1986). Hörertypologien. Zur Psychologie des jugendlichen Musikgeschmacks. Regensburg: Bosse.

[6] Besseler, H. (1926). Grundfragen des musikalischen Hörens. Jahrbuch der Musikbibliothek, 32, 35-52.

[7] Bourdieu, I. (1982). Die feinen Unterschiede. Frankfurt am Main: Suhrkamp.

[8] Brosch, T. et al. (2010). The perception and categorisation of emotional stimuli: A review. Cognition Emotion, 24(3), 377–400. https://doi.org/10.1080/02699930902975754.

[9] Brömse, P. and Kötter, E. (1971). Zur Musikrezeption Jugendlicher: Eine psychometrische Untersuchung. Mainz: Schott.

[10] Brüggemann, S. (2017). Das Musikkonsumverhalten von Streaming-Nutzern im Kontext der Medien- und Produktkonfigurationen (Master Thesis). München: Ludwig-Maximilians-Universität.

[11] Cespedes-Guevara, J. and Eerola, T. (2018). Music communicates affects, not basic emotions: A constructionist account of attribution of emotional meanings to music. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.00215.

[12] Charron, J. (2017). Music Audiences 3.0: Concert-Goers' Psychological Motivations at the Dawn of Virtual Reality. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.00800.

[13] Clarke, E. and Cook, N. (2004). Empirical musicology: Aims, methods, prospects. Oxford: Oxford University Press.

[14] de la Motte-Haber, H., Kopiez, R., and Röttger, G. (1996). Handbuch der Musikpsychologie. Laaber: Laaber.

[15] de la Motte-Haber, H. and Neuhoff, H. (2007). Musiksoziologie. Laaber: Laaber.

[16] Dirks, T. (2016). Digitalisierung für den Menschen - Jederzeit, an jedem Ort. Retrieved November 22, 2019, from https://www.telefonica.de/2016/10/ceo-thorsten-dirks-digitalisierung-fuer-den-menschen-jederzeit-an-jedem-ort/.

[17] Dollase, R. (1986). Musikpräferenzen und Musikgeschmack Jugendlicher. In D. Baacke (Ed.), Handbuch Jugend und Musik (pp. 341-368). Opladen: Leske Budrich.

[18] Ericsson Consumerlab (2017). TV and Media 2017. A consumer-driven future of media. Retrieved from https://www.ericsson.com/en/reports-and-papers/consumerlab/reports/tv-and-media-2017.

[19] Farnsworth, P. (1958). The social psychology of music. New York: Dryden Press.

[20] Fuller, Z. (2018). Do Playlists Make the Power Law in Music Even More Extreme? Retrieved January 3, 2020, from https://www.midiaresearch.com/blog/do-playlists-make-the-power-law-in-music-even-more-extreme/.

[21] Gembris, H. (1999). 100 Jahre musikalische Rezeptionsforschung: Ein Rückblick in die Zukunft. Musikpsychologie, 14, 24-41.

[22] Gracenote (2019). Global music data - Gracenote. Retrieved November 22, 2019, from https://www.gracenote.com/music/global-music-data/.

[23] Grekow, J. (2018). Audio features dedicated to the detection and tracking of arousal and valence in musical compositions. Journal of Information and Telecommunication, 2(3), pp. 322–333. https://doi.org/10.1080/24751839.2018.1463749.

[24] Herrmann-Sinai, S. (2009). Sounds without the mind? Versuch einer Bestimmung des Klangbegriffs. Deutsche Zeitschrift Für Philosophie, 57(6), 885–906. https://doi.org/10.1524/dzph.2009.0074.

[25] Herzog, M. et al. (2017). Predicting musical meaning in audio branding scenarios. Berlin: Universität Berlin.

[26] Holt, F. (2010). The economy of live music in the digital age. European Journal of Cultural Studies, 13(2), 243–261. https://doi.org/10.1177/1367549409352277.

[27] Horkheimer, M. and Adorno, T. W. (1981). Dialektik der Aufklärung: Philosophische Fragmente. Frankfurt am Main: Suhrkamp.

[28] Int. Federation of the Phonographic Industry (2018). Global music report 2018 - Annual state of the industry. Retrieved from https://www.ifpi.org/downloads/GMR2018.pdf.

[29] Johnson, L. (2018). Spotify's latest feature tells you even more about music, podcasts and audio books. Retrieved November 22, 2019, from https://www.techradar.com/news/spotifys-latest-feature-tells-you-even-more-about-music-podcasts-and-audio-books.

[30] Joven, J. (2018). The rise of contextual playlists. Chartmetric [Blog post]. Retrieved November 22, 2019, from https://blog.chartmetric.io/spotify-the-rise-of-the-contextual-playlist-c6f2c26900f4.

[31] Juslin, P. (2013). What does music express? Basic emotions and beyond. Frontiers in Psychology, 4. https://doi.org/10.3389/fpsyg.2013.00596.

[32] Kachkach, A. (2016). Analyzing user behaviour and sentiment in music streaming services (Master Thesis). Stockholm: KHT Royal Institute of Technology.

[33] Kassabian, A. (2013). Ubiquitous listening: Affect, attention, and distributed subjectivity. Berkley: University of California Press.

[34] Kloppenburg, J. (2009). Musikpräferenzen. In H. De la Motte-Haber G. Rötter (Eds.), Musikpsychologie (2nd ed., pp. 357–393). Laaber: Laaber.

[35] Knees, P. and Schedl, M. (2016). Music Similarity and Retrieval: An Introduction to Audio- and Web-based Strategies. Berlin: Springer.

[36] Kornhuber, H. and Deecke, L. (2010). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. Pflügers Archiv Für Die Gesamte Physiologie Des Menschen Und Der Tiere, 284(1), 1–17. https://doi.org/10.1007/bf00412364.

[37] Krüger, H. (1999). Zwischen Lachen und Weinen: Das Spektrum menschlicher Phänomene. Berlin: Akademie Verlag.

[38] Lamere, P. (2014). The skip. Retrieved November 22, 2019, from https://musicmachinery.com/2014/05/02/the-skip/.

[39] Larsen, J. (2009). The music industry and digital music: Disruptive technology and the value network effects on industry incumbents. Copenhagen: Copenhagen Business School.

[40] Mangold, R. et al. (2004). Lehrbuch der Medienpsychologie. Göttingen: Hogrefe.

[41] Mark, D. (1998). Wem gehört der Konzertsaal? Das Wiener Orchesterrepertoire im internationalen Vergleich. Zur Frage des musikalischen Geschmacks bei John H. Mueller. Wien: Guthmann-Peterson.

[42] Miles, S. (2018). The emergence of contemporary consumer culture. In O. Kravets, P. Maclaran, S. Miles, A. Venkatesh (Eds.), The SAGE handbook of consumer culture (pp. 13–26). Los Angeles: SAGE Publications.

[43] Müller, R. et al. (2002). Wozu Jugendliche Musik und Medien gebrauchen: Jugendliche Identiät und musikalische und mediale Geschmacksbildung. Weinheim: Juventa.

[44] Müller-Freienfels, R. (1936). Psychologie der Musik (pp. 35-52). Berlin-Lichtenfelde: Vieweg Teubner.

[45] Münch, T. (1998). 24 Stunden in 3 Minuten? Computergestützte Musikprogrammerstellung im Radio der 90er Jahre. In B. Endres, N. Knolle (Eds.), KlangArt-Kongreß 1995 (pp. 399–414). Osnabrück: Universitätsverlag Rasch.

[46] Münchner Kreis (2013). Innovationsfelder der digitalen Welt: Bedürfnisse von übermorgen. München: Münchner Kreis.

[47] Münchner Kreis (2016). Neue Produkte in der digitalen Welt. München: Übernationale Vereinigung für Kommunikationsforschung e.V. and Münchner Kreis.

[48] Nawaz, R. et al. (2018). Acoustic feature extraction from music songs to predict emotions using neural networks. 2018 2nd International Conference on BioSignal Analysis, Processing and Systems (ICBAPS), 166–170. https://doi.org/10.1109/icbaps.2018.8527414.

[49] Nielsen Holdings (2017). U.S. music 360: 2017 report highlights. Retrieved from https://www.nielsen.com/us/en/insights/report/2017/music-360-2017-highlights/.

[50] North, A. and Hargreaves, D. (1997). Experimental aesthetics and everyday music listening. In D. Hargreaves, A. North (Eds.), The social psychology of music (pp. 84–103). Oxford: Oxford University Press.

[51] Nylund-Hagen, A. (2015). The playlist experience: Personal playlists in music streaming services. Popular Music and Society, 38(5), 625–645. https://doi.org/10.1080/03007766.2015.1021174.

[52] Nylund-Hagen, A. (2016). The metaphors we stream by: Making sense of music streaming. First Monday, 21(3). https://doi.org/10.5210/fm.v0i0.6005.

[53] Oehmichen, E. and Ridder, C. M. (2003). Die MedienNutzerTypologie: Ein neuer Ansatz der Publikumsanalyse. Baden Baden: Nomos.

[54] Pessoa, L. (2008). On the relationship between emotion and cognition. Nature Reviews Neuroscience, 9(2), 148–158. https://doi.org/10.1038/nrn2317.

[55] Plessner, H. (1980). Gesammelte Schriften: Anthropologie der Sinne. Frankfurt am Main: Suhrkamp.

[56] Posner, J. et al. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and Psychopathology, 17(03), 715–734. https://doi.org/10.1017/s0954579405050340.

[57] Prey, R. (2017). Nothing personal: Algorithmic individuation on music streaming platforms. Media, Culture Society, 40(7), 1086–1100. https://doi.org/10.1177/0163443717745147.

[58] Pöppel, E. and Welker, L. (2009). Ein anderer Ton. Das Hofer Modell. München: HWX Ludwig-Maximilians-Universität.

[59] Ross, U. (2010). Entspannung: Neuropsychobiologische Aspekte einer vernachlässigten Selbstverständlichkeit. Swiss Journal of Integrative Medicine, 22(2), 100–113. https://doi.org/10.1159/000284116.

[60] Schröger, E. (2005). The mismatch negativity as a tool to study auditory processing. Acta Acustica, 91(3), 490-501.

[61] Schulze, G. (1992). Die Erlebnisgesellschaft: Kultursoziologie der Gegenwart. Frankfurt am Main: Campus.

[62] Seemann, J. (2012). Hybrid insights: Where the quantitative meets the qualitative. Retrieved November 23, 2019, from https://store.hbr.org/product/hybrid-insights-where-the-quantitative-meets-the-qualitative/rot182?sku=ROT182-PDF-ENG.

[63] Sinus Sociovision (2007). Die Sinus-Milieus in Deutschland 2007. Retrieved November 22, 2019, from https://www.sinus-institut.de/sinus-loesungen/sinus-milieus-deutschland/.

[64] Sloboda, J. (2012). Choosing to hear music: Motivation, process and effect. In S. Hallam, I. Cross, M. Thaut (Eds.), Oxford Handbook of Music Psychology (pp. 431–490). Oxford: Oxford University Press.

[65] Smudits, A. (2007). Wandlungsprozesse der Musikkultur. In H. De la Motte-Haber H. Neuhoff (Eds.), Musiksoziologie (pp. 111–145). Laaber: Laaber.

[66] Spotify (2018). Our Spotify cheat sheet: 4 ways to find your next favorite song. Retrieved November 23, 2019, from https://newsroom.spotify.com/2018-11-02/our-spotify-cheat-sheet-4-ways-to-find-your-next-favorite-song/.

[67] Spotify (2019a). The basics - Spotify. Retrieved November 23, 2019, from https://support.spotify.com/us/using_spotify/the_basics/.

[68] Spotify (2019b). Get audio features for a track: Spotify for developers. Retrieved November 23, 2019, from https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/.

[69] Stiftung für Zukunftsfragen (2018). Die Zukunft der Medien: Die Zukunft der Medien: Alte und „neue“ Medien verschmelzen nur langsam. Retrieved November 23, 2019, from https://www.stiftungfuerzukunftsfragen.de/forschung/forschungsthemen/die-zukunft-der-medien/.

[70] Stocks, E. (2017). Music consumption in the era of smart speakers. Retrieved November 23, 2019, from https://medium.com/@elliotjaystocks/music-consumption-in-the-era-of-smart-speakers-b88d04a18746.

[71] Treseder, D. (2018). What's the next social network? Think music. Retrieved November 22, 2019, from https://www.adweek.com/digital/whats-the-next-social-network-think-music/.

[72] University of Southern California (2018a). Organizing your social sciences research paper: Qualitative methods. Retrieved November 23, 2019, from https://libguides.usc.edu/writingguide/qualitative.

[73] University of Southern California (2018b). Organizing your social sciences research paper: Quantitative methods. Retrieved November 23, 2019, from https://libguides.usc.edu/writingguide/quantitative.

[74] Vaitl, D. and Petermann, F. (2000). Handbuch der Entspannungsverfahren: Grundlagen und Methoden (2nd ed.). Weinheim: Psychologie-Verlags-Union.

[75] van den Hoven, J. (2015). Analyzing Spotify data: Exploring the possibilities of user data from a scientific and business perspective (Research Thesis). Amsterdam: VU Amsterdam.

[76] Vargo, S. and Lusch, R. (2008). Service-dominant logic: continuing the evolution. Journal of the Academy of Marketing Science, 36(1), 1–10. https://doi.org/10.1007/s11747-007-0069-6.

[77] Wikström, P. (2012). A typology of music distribution models. International Journal of Music Business Research, 1, 7–20. Retrieved from https://musicbusinessresearch.wordpress.com/international-journal-of-music-business-research-ijmbr/.

[78] Winterhoff-Spurk, P. (1999). Medienpsychologie: Eine Einführung. Köln: Böhlau.