The London School of Economics and Political Science

MATCHMAKERS OR *TASTE*MAKERS?

Platformization of Cultural Intermediation
&
Social Media's Engines for 'Making up Taste'

Zeynep Kaparoglu

A thesis submitted to the Department of Management, Information Systems and Innovation Group, of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London

January 2018

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 81,914 words (including footnotes but excluding references).

Table of Contents

Abstract

There are long-standing practices and processes that have traditionally mediated between the processes of production and consumption of cultural content. The prominent instances of these are: *curating* content by identifying and selecting cultural content in order to promote to a particular set of audiences; *measuring* audience behaviours to construct knowledge about their tastes; and *guiding* audiences through recommendations from cultural experts. These cultural intermediation processes are currently being transformed, and social media platforms play important roles in this transformation. However, their role is often attributed to the work of users and/or recommendation algorithms. Thus, the processes through which data about users' taste are aggregated and made ready for algorithmic processing are largely neglected. This study takes this problematic as an important gap in our understanding of social media platforms' role in the transformation of cultural intermediation. To address this gap, the notion of *platformization* is used as a theoretical lens to examine the role of users and algorithms as part of social media's distinct data-based sociotechnical configuration, which is built on the so-called 'platform-logic'. Based on a set of conceptual ideas and the findings derived through a single case study on a music discovery platform, this thesis developed a framework to explain 'platformization of cultural intermediation'. This framework outlines how curation, guidance, and measurement processes are 'plat-formed' in the course of development and optimisation of a social media platform. This is the main contribution of the thesis. The study also contributes to the literature by developing the concept of social media's engines for 'making up taste'. This concept illuminates how social media operate as sociotechnical cultural intermediaries and participates in tastemaking in ways that acquire legitimacy from the long-standing trust in the objectivity of classification, quantification, and measurement processes.

List of Figures

List of Tables

The following images have been redacted as the copyright is owned by another organisation.

Page 104
Page 105
Page 107
Page 108
Page 109
Page 110
Page 112
Page 113
Page 114
Page 120
Page 122
Page 123
Page 124
Page 125
Page 126
Page 127
Page 129
Page 130
Page 134
Page 135
Page 136
Page 137
Page 138
Page 149
Page 160
Page 241
Page 242-247

1. Introduction

Transformations enabled by social media have implications that are far-reaching. They have consequently inspired research from a range of disciplines, including Information Systems (IS). Cross-functional and cross-disciplinary aspects of the research on social media assume a central role for IS researchers in this domain (Aral, Dellarocas and Godes, 2013). As a theoretical project in the cross-territory between media and information systems, this dissertation aims to contribute to this domain. It is primarily concerned with cultural transformations enabled by social media as *platform-based* businesses and the main purpose of the thesis is to explain how social media platforms are transforming cultural intermediation. This locates the thesis within the area of the interdisciplinary IS field known as Social Studies of Information and Communication Technologies (SSICTs). This stream of research is particularly concerned with the social and cultural transformations enabled by information and communication technologies (ICTs) as they become increasingly intertwined with numerous aspects of everyday living.

This chapter first outlines the motivation to research the cultural transformations enabled by social media platforms, with a particular focus on how they alter conditions of tastemaking. This is followed by a delineation of the methodological approach chosen to conduct this research. The chapter concludes with an outline of the thesis.

1.1. Motivation and Object of Research

1.1.1. Cultural Intermediation and Tastemaking

Thirty years ago, the local radio DJ finding new and interesting music for her radio audience used to be accepted as a legitimate tastemaker. Similarly, the owner of a record store was a trusted source to get personalized music recommendations matching audience tastes. Critics and reviewers guided tastes of audiences, and their tacit knowledge and ability to read audience tastes rendered them experts in specific cultural domains. This legitimised their role as tastemakers. Thus, they long served as reliable sources informing us about what is new, what is good, or what is trending, and accordingly guided tastes. Likewise, media and cultural industries have employed audience research techniques and measurement technologies in order to know more about audience and guide their tastes. Findings of audience measurement were used to inform programming or production decisions. As a result, what was produced,

framed, and made available for audiences to consume was predetermined. These gatekeepers collectively played key roles in the identification, selection, and promotion of particular cultural content to be made available to audiences for cultural consumption. In doing so, they have long been influential on taste formation.

Pierre Bourdieu (1984) used the term 'cultural intermediary' in his seminal work, *Distinction*, to point out the tastemaking role of a specific group of actors affecting the circulation of cultural goods. Cultural intermediaries are defined as 'professional tastemakers' (Smith Maguire and Matthew, 2014) shaping cultural tastes, because they are involved in the presentation and representation of cultural content and framing how others encounter it. Their tastemaking ability is cemented as they gain legitimacy from their expertise in a given field or based on their demonstrated ability to guide tastes. In a similar vein, contemporary sociotechnical configurations, such as social media, play significant roles in framing people's online encounters with cultural content. From this perspective, this study investigates how social media platforms transform the cultural intermediation process and participate in tastemaking.

One of the fundamental assumptions upon which this research project is built is that our experience and understanding of taste do not emerge 'naturally' from our sensory engagement with the material world around us (Wright, 2015). Rather, the processes in which the cultural tastes are formed are considered as being traditionally 'conditioned.' There are various factors that are influential on taste formation. These can be broadly subsumed under two main categories: the *consumption* side and the *production* side. The fundamental social dimensions – of class, gender, ethnicity, age, and so on – are known to be substantially influential on the formation and expressions of taste (ibid.). Investigations of these dimensions are usually concerned with 'how *exactly* audience tastes are formed,' and 'how audiences *like* what they *like'*. Such factors concerning the consumption side are not the concerns of this research project. That is to say, this research does not investigate taste formation by studying user populations. Instead, processes comprising the production side are the objects of interest here.

Processes related to the production side *precede* cultural consumption, and *condition* how individuals (consumers and/or audiences) encounter and consume cultural content. This study is explicitly concerned with investigating the processes *mediating* between the production and consumption of cultural content. These are conceptualised here as cultural intermediation

processes, the common defining attribute of which is considered to be 'tastemaking'. The main focus of the study is to understand the ways in which these processes are being transformed by social media platforms. Toward this end, the thesis conceptualises cultural intermediation as comprised of three key processes that mediate between the production and consumption of cultural content: i) curation, ii) guidance, and iii) measurement. These are taken as the key instances of the infrastructural processes underpinning symbolic production, which includes, but is not limited to: people, processes, practices, norms, assumptions, techniques, technologies in the media and culture industries. Collectively, they are assumed to play key roles in determining and framing what content audiences encounter and how they do so, in subtle yet highly significant ways. This, in turn, affects how audiences interact with and experience cultural content, and is thereby assumed to shape their cultural tastes.

1.1.2. Social Media

There is an increasing recognition of how social media transform various social and economic phenomena, including cultural intermediation. In the current state of the literature, this transformation is to a large extent studied at the user-interface level activities through which ordinary individuals find opportunities for attending curation, guidance, and gatekeeping. On the other hand, attributing significance to the role of algorithms mystifies the role of the platform in these processes. The literature on social media acknowledges how social platforms shape user participation through their prescriptive design and complex data operations. This recognition is an important pillar of social media research (Aral et al., 2013). However, there is a shortage in studies unpacking the relationship between the frontend and backend of social media to investigate their role in the transformation of cultural intermediation. This constitutes a gap in our understanding of the cultural implications of recent sociotechnical developments like social media platforms and the increasingly pervasive platform-based business models. To fill this gap, the aim of this dissertation is to unpack the sociotechnical configuration of social media and lay bare the relationship between their frontend and backend. In doing so, the main focus will be on deconstructing how social media operate as matchmakers; that is, how they create data representations of individuals to provide them with personalised recommendations, which are assumed to '*match'* their '*taste'*.

'Matchmaker' (Evans and Schmalensee, 2016) is a term widely applied to platform-based businesses, because the fundamental logic behind them is matchmaking; that is, bringing together different participants, connecting them in a central infrastructure, and facilitating

interactions between them. The fundamental role of a platform in this setting is to identify relevance (Gillespie, 2014) among participants and the content and thus consummate matches among them (Parker et al., 2016). This matchmaking capacity renders social media a novel form of cultural intermediary. Matchmaking processes necessitate *curation*, which is the identification and selection of particular units for exchange between users based on the taste profiles that the platform system constructs for them. This requires the system to capture users' taste by *measuring* their online behaviours to turn them into computable data for constructing their digital taste profiles. The items identified to be relevant for these digital taste profiles are then promoted to the users through personalised recommendations, thereby *guiding* them towards particular content and implicitly moving them away from others.

All of these operations and processes can be traced back to traditional cultural intermediation processes. Social media platforms transpose these processes to digital context and introduce some unique qualities. In this regard, the main thesis of this dissertation is that social media platforms operate as sociotechnical cultural intermediaries. The main focus is to explain the relationship between the frontend and backend of social media in ways that illuminate how social media operate as matchmakers. Doing so will pave the way for discussing the core assumption underpinning this thesis: Social media platforms are *not* merely matchmakers; in the course of their matchmaking operations they also operate as tastemakers.

1.2. Problem Domain and Epistemological Stance

The main problematic of this study is the importance attributed to the work of users and/or algorithms to explain how cultural intermediation processes are altered by social media, thereby overlooking the role of platform design and data operations of social media in this transformation. This creates a gap in the literature concerned with social media's role in current cultural transformations. To address this gap, this research adopts the stance taken by Wegner (1997) in his seminal work on the significance of interactive data over algorithms. In addition it adopts the stance taken by the scholars who critically engage with the assumptions regarding the non-neutrality of user activities on social media and the ways in which social data are produced by these platforms (Alaimo 2014; Alaimo and Kallinikos, 2016, 2017; Bucher, 2012; Couldry and Kallinikos, 2017; Langlois and Elmer, 2013; van Dijck, 2013, 2014). From this perspective, this thesis unpacks the processes in which data representations of the individuals' tastes are produced and acted upon in the matchmaking operations of

social media. The purpose behind this is to shed light on the processes that make data ready for algorithms, thereby placing users and algorithms within the larger data arrangements of which they are a part. Thus, the thesis aims to explain how social media attend to conditioning of taste formation in unique ways enabled by its distinct sociotechnical configuration.

This thesis starts with the assumption that the processes involved in conditioning taste formation are socially constructed. However rather than merely focusing on their social constructedness the main concern here is to understand what *kind* of a conditioning is taking place. This point connects to the epistemological stance adopted in this study, which highlights, following Hacking (1999), the importance of understanding *construction of kinds*. This study does not assume social and cultural outcomes, such as taste formation, to be produced mechanically in ways dictated by technology. Individuals certainly have different forms of agency and do not blindly follow only what is offered by technology. To what extend technology structures individuals' interaction with and experience of cultural artefacts is subject to their opting to use the technology. However, as emphasised throughout this dissertation, the frontend design and backend data operations of social media are intertwined in highly complex ways, which are not easily accessible to and/or comprehensible by ordinary individuals. The rationale underpinning such operations are not accountable, even to those who have sufficient knowledge and skills to understand data-based operations. These procedures are protected under intellectual property rights because they are the most valuable assets of the companies running these platforms. In this regard, users of social media often act upon what they see on the interface level without knowing *how* they see it, as the logic behind the production of personalized suggestions is not visible to them. In turn, they feed the system with their subsequent actions, i.e., by listening, watching, 'liking', or skipping what is shown to them as personalised recommendations that are 'tailored' to their taste. Thus, what they see based on the platform's assumption regarding 'what is relevant to their taste' become prone to shape 'what they actually see and consume next', in a cyclical manner.

In this regard, this study argues that social media transform cultural intermediation and recondition formation of cultural taste in highly subtle ways. It is important to understand what kind of a conditioning this is without taking a technologically determinist stance. I do acknowledge that the human-computer interaction that takes place at the interface-level varies across contexts and situations and individuals have agency to negotiate what is offered

by technology (Kallinikos, 2004). However, overlooking what is happening beneath the interface level runs the risk of leaving our knowledge rather incomplete. To eliminate that risk, this study unpacks the sociotechnical configuration of social media by deconstructing the relationship between the frontend and backend to explain how social media recondition taste formation.

1.3. Research Project

The empirical investigation for the aforementioned research domain is based on a single explanatory case study. Empirical evidence is collected from a typical representative of a social media platform operating in the domain of music discovery. Based on this case study, a theoretical framework is developed to explain how social media transform cultural intermediation. The process of theory building starts with an initial literature review, the identification of the explanatory concepts, and the development of a preliminary framework. This is followed by gradual refinement of the ideas put forward in this initial model in light of the empirical findings. Thus the model takes its final form. Resultant explanation paves the way for opening up a critical discussion on the implications of the matchmaking operations of social media as contemporary form of cultural intermediaries.

1.3.1. Theoretical Framework and Research Questions

Several streams of research inform the conceptual framework devised to guide the analysis of the findings, namely the literature on platforms, social media, cultural intermediation, audiencemaking and the literature concerned with the classification, quantification and measurement. Drawing on the previous works in these streams of research, *Platformization* and *Cultural Intermediation* are identified as the main explanatory concepts. The main argument is that social media platforms are designed and developed based on a distinctive rationale, and their functioning is sustained by their unique data-based operations. It is this unique relationship between design and data operations that underpins their matchmaking operations. Matchmaking enables them to operate as intermediaries facilitating interactions among their users, who are the producers and consumers of the digital cultural content circulated on these platforms. In other words, users encounter with and experience content in accordance with the logic underlying the platform's matchmaking operations. It is important to deconstruct this logic to explain how they alter traditional cultural intermediation processes involved in shaping cultural tastes. In this regard, the main research question posed in this

study is: *How do social media platforms operate as matchmakers, and, in doing so, how do they transform cultural intermediation?* To address this question, a conceptual framework is devised to explain what can be referred to as the 'platformization of cultural intermediation' by social media.

1.3.2. Case Study

The 'how' questions put forward in this study require an in-depth investigation of the phenomenon under scrutiny through prolonged immersion into its real-life context (Yin, 2003). This especially holds true when it comes to challenges involved in investigating complex sociotechnical arrangements like social media (Couldry and Kallinikos, 2017). Case study is thus chosen as the research strategy. It enables a researcher to build a detailed understanding of the research domain by allowing her to collect data from multiple sources of evidence. Thus the researcher becomes able to give a thorough explanation of the researched phenomenon.

The domain of music is selected as the taste case to investigate the arguments put forward in this study regarding the transformation of cultural intermediation processes, and, by extension, alteration of the tastemaking. This choice is due to the frequency, intimacy, and ubiquity of music listening (Kassabian, 2013), and hence the significance of the role of cultural intermediation in music discovery and listeners' interaction and experience with music. For this reason, a social media platform specialised is music discovery is chosen for empirical investigation. The chosen case is a typical representative of social media platforms operating as matchmakers.

The case study finds that the distinct design logic and the data operations of social media are closely linked, and their relationship plays a significant role in how social media perform matchmaking among participants. Thus they operate as cultural intermediaries attending to curation, guidance, and measurement in unique ways. This validates the propositions made based on the initial framework devised to guide this study. These preliminary ideas are expanded in light of the analysis of the empirical findings. This paved the way for a gradual development of the theory through a cyclical process continuously informed by theory and data. Consequently, the final explanatory framework outlines the *development* and *optimisation* of a social media platform as the key aspects of the platformization of cultural intermediation.

1.3.3. Contributions

The main contribution of this study is the development of an explanatory framework to theorise platformization and cultural intermediation. This framework outlines how curation, guidance, and measurement processes are *internalised* by a social media platform during its *development* and *optimisation*. It also highlights how these processes merge, get layered and shared out by the social media users and the platform system in the course of being internalised by the platform. As part of this theorising, social media platforms are conceptualised as 'sociotechnical cultural intermediaries' and their distinct aspects are discussed. As part of this discussion, a new concept is proposed: social media's engines for 'making up taste'. This concept is developed to delineate the implications of the data means social media use to measure people's taste. These means cyclically operate and lead to a mutual calibration of platform functionalities and user behaviour and are thereby prone to operate in a performative fashion. In this regard, this thesis contributes to a specific IS literature known as Social Studies of Information and Communication Technologies (SSICT) by highlighting some the cultural transformations enabled by social media platforms in ways that open up a critical discussion regarding the taken for granted aspects of social media. In doing so, it also contributes to the literature on social media, platforms, and cultural intermediation, as well as a wider branch of the literature concerned with the implications of classification and measurement.

1.4. Outline of the Thesis

Following this Introduction, Chapter 2 provides a review of the literature on cultural intermediation and social media and highlights how their conceptualisation evolved over time. This chapter sets the context and the background of the study. It draws on the notion of cultural intermediation to unveil prominent instances of the processes mediating between the production and consumption of cultural content to better understand how these processes are being transformed by social media platforms. Against this backdrop, the chapter reviews current views on how social media are involved in the transformation of the cultural intermediation process. It does so by looking at curation, guidance, and measurement in social media as the key instances of cultural intermediation.

Drawing on the existing views this review highlights that the predominant approach to understand how social media transform cultural intermediation is attributing the novelty to the work of users and/or the algorithms. Current views suggest that the mediators of the past

(i.e. gatekeepers, curators, and reviewers, media companies etc.) that 'collectively dictated' (Bhaskar 2016a) what content should be consumed, is now being replaced by a complex mixture of algorithms and user-curations. These views are problematized on the grounds of neglecting the role of data in explaining the transformation of cultural intermediation. Thus this chapter sets the background of the study, introduces its main problematic, and puts forward its main argument as an emphasis of the significance of data and data production in social media over user practices and algorithms.

Chapter 3 reviews the key insights from the literature on platforms and social media and devises a preliminary framework to guide the empirical investigation. Here I discuss the main concepts guiding the case study and delineate the relationship between them. I use the concepts of 'platformization' and 'cultural intermediation' and outline how a social media platform internalises cultural intermediation process through the distinct relationship between its frontend and backend, which underpins its matchmaking operations. Finally, in light of this framewor, I refine the main research question and develop sub-questions to be operationalised during the case study.

In Chapter 4, I explain the research strategy used to answer the research question. I first explain the epistemological stance of the study and reflect on the methodological decisions. This is followed by a discussion of the case study design, with corpus construction as the chosen data collection technique, and the interpretation and generalisation of findings. Here I describe how this study is driven by a preliminary conceptual lens and aims to develop a theory. Accordingly, I explain how the study is based on a cyclical process of data collection, analysis, and further data elicitation for subsequent rounds of analysis. Thus I discuss in this chapter how the resultant theory is built gradually through an iterative analysis process that is informed by both theory and empirical evidence.

Chapter 5 presents the case study narrative. The chapter first contextualises the case and introduces the company behind the social media platform chosen for the empirical study. Then the case study narrative is presented. This section is divided into two subsections, wherein the presentation of the empirical narrative is organised based on the description of the *platform* and the *process* in which it is developed and evolved. First, the basic structure of the three different versions of the platform is depicted in a sequential manner. This is followed by the description of the process in which the platform configuration has been changed three times. This section highlights the reasons and logic behind the changes made

in the platform and describes its evolution over time.

Chapter 6 presents the findings of the case study. Discussion of findings is organised around two main themes: i) major challenges and concerns, and ii) tools and techniques used to address those challenges. The first section describes the main issues faced by the developers during the development of the platform, which paved the way for changing the platform configuration three times. The second section outlines the strategies, as well as the tools and techniques, used to address these challenges.

Chapter 7 applies the conceptual framework to the case study findings in order to answer the research questions. This chapter finds that the unique ways in which a social media platform is developed is highly influential in transforming cultural intermediation. Analysis of empirical evidence in this chapter reveals that several concerns underpin the process of developing a social media platform and these concerns significantly shape how a platform internalises curation, guidance, and measurement. These findings validate the initial ideas regarding the role of social media's distinct design and complex data operations in altering cultural intermediation processes. However, findings presented in Chapter 6 reveal that there are additional factors that should be taken into account in order to further develop the framework outlining this transformation. These additional themes are identified as elements of platform optimisation and are discussed in the second part of the chapter. Thus, in this final stage of theory development, an explanatory framework is developed to delineate the platformization of cultural intermediation.

Chapter 8 discusses the explanatory framework in ways that highlight how this research contributes to the different streams of research that informed this study. This paves the way for building on and extending previous works and conceptualising social media as sociotechnical cultural intermediaries. Implications of this are discussed in relation to the notions of tastemaking in light of wider debates in the literature concerned with the performative implications of the processes of categorisation, quantification, and measurement. Thus a new concept is proposed: social media's engines for 'making up taste'.

Chapter 9 is the conclusion chapter, in which an overview of the thesis is presented, contributions and limitations are discussed, and possibilities for future research are outlined. I conclude by highlighting how social media should be considered as contemporary cultural

intermediaries, which not only operate as matchmakers but also as sociotechnical tastemakers legitimatised by the long-standing trust in the objectivity of the numbers.

2. Literature Review

2.1. Introduction

This chapter reviews and synthesises the literature on cultural intermediation and social media. Then it highlights the views on how social media transform cultural intermediation. The outline of the chapter is as follows:

I first provide a brief review of the literature on *cultural intermediaries,* outlining their defining attribute as 'tastemakers'. Following this, cultural intermediation is cast under a new light, one that suggests considering it as a *process* rather than as being restricted to the work of people and certain occupational groups. This section serves as the foundation upon which the notion of *cultural intermediation* is conceptualised in this study. Synthesizing previous works and drawing on the key insights from the literature, the chapter then conceptualises *curation*, *guidance* and *measurement* as the key instances of the cultural intermediation process. This section's concluding section problematizes the shortage of empirical studies scrutinising the sociotechnical configuration of how these processes are being altered by sociotechnical developments, such as social media platforms. Thus I move to review the literature on social media. I outline its theoretical evolution, from the conceptualization of social media as networking sites to platforms infrastructuring mundane social and cultural interactions. This part is followed by a review of previous works concerned with understanding how social media transform cultural intermediation. In the conclusion, I problematize the overemphasis in this literature on the frontend user performances and backend algorithmic operations in ways that lead to neglecting the relationship between the frontend and backend of social media and the role data plays in this relationship.

2.2. Literature on Cultural Intermediation

This section sets the context and the background of the study. It draws on the notion of cultural intermediation to unveil prominent instances of the processes mediating between the production and consumption of cultural content. The inquiry starts with introducing the concept of a cultural intermediary and a brief review of the previous work on it. It then proceeds to conceptualise cultural intermediation as a process and outlines the main activities of which it is comprised.

It is worthwhile to start with some definitions for conceptual clarity. First of all, this section is particularly interested in identifying the processes in between the production and

consumption of cultural goods. By cultural goods, I mainly refer to media and cultural content, such as music, books, and film, subsumed under the conceptual category: 'cultural content'. In the same vein, I refer to the consumer of cultural content as 'audience'. Thus, hereinafter, I use the terms 'cultural content' and 'audience' to refer to these broader categories in order to avoid conceptual confusion.

2.2.1. The Concept of a Cultural Intermediary

Sociology of cultural consumption is a specific branch of literature concerned with symbolic goods and how they acquire meaning. In this literature, Pierre Bourdieu famously uses the concept of a 'cultural intermediary' in his seminal work Distinction. Bourdieu (1984 p.1011) defines cultural intermediaries as a group of tastemakers who focus on the production of 'consuming' tastes and dispositions. They are described as a group of professionals that are involved in the 'gentle manipulation' of tastes for cultural products (Bourdieu 1984, p. p. 365). Their work connects culture and economy as they strategically involve in the production and promotion of consumption by constructing legitimacy and producing symbolic value for cultural goods.

From this perspective, cultural intermediaries are conceptualized as strategically located actors that play important roles in the space that mediates between the producers and audiences of cultural content (Smith-Maguire and Matthew, 2014). They are involved in the framing, qualification and circulation of symbolic goods and services and hence significantly shape how these are experienced. The most prominent feature of cultural intermediaries is argued to be their participation in tastemaking. They accomplish this through the legitimacy and authority they acquire in the eye of the public and are key figures within cultural industries based their expertise in a particular field (Smith-Maguire, 2014).

In the following section, I explain this tastemaking function as a defining attribute of cultural intermediaries as legitimized through the expertise and credibility they acquire in a specific domain.

i. Professional Tastemakers

Cultural intermediaries are defined by their work as tastemakers (Smith-Maguire and Matthew, 2014, p.20). This is intertwined with Bourdieu's understanding of how taste operates as 'a match-maker' between people and things (Bourdieu, 1984 p.243). According to

Smith-Maguire and Matthew (2014), cultural intermediaries cannot enforce desires or purchases; rather, they create the conditions for consumers to identify their tastes in goods. In doing so, they build 'elective affinities' among cultural content and consumer tastes. Smith-Maguire and Matthew (2014) illustrate this by drawing on how an advertising executive, as a prominent cultural intermediary, explains her job: "In my business, we're constantly classifying people. There are social classes, castes, and it's a matter of fitting a product to the right caste" (Smith-Maguire and Matthew, 2014 p.20). This indicates that, by imposing symbolic meanings, cultural intermediaries frame cultural content in ways that appear to go well with the audience's tastes. A set of complementary actors (i.e. media) is considered to support the ways in which cultural intermediaries channel consumers' perceptions and preferences. Cultural intermediaries gain authority for tastemaking through their ability to perform a matchmaking role between consumers and cultural goods. This is legitimised based on their tacit knowledge and expertise in a given field. In this regard, cultural intermediaries are not simply tastemakers; they are considered 'professional tastemakers' (Bourdieu, 1990 p. 96), who have the authority to legitimise what should be consumed. These professional tastemakers are considered to create the 'conditions' for the identification of tastes for certain cultural goods (Webster et al., 2016).

The concept is first used to describe a group of occupational actors involved in tastemaking between cultural production and consumption, such as the producers of cultural programs in radio and television, advertising and marketing creatives, critics, museum curators and magazine reviewers (Smith-Maguire and Matthew, 2014; Webster et al., 2016). The range of people functioning as tastemakers has extended since then (Ashton and Couzins, 2015) to include those who work in supporting roles and/or promotional activities, such as fostering certain 'lifestyle' choices (Featherstone, 1991). The term is thus extensively used to study the production of meaning and value in different industrial contexts and value chains to highlight implications of the symbolic and material practices involved in the promotion of consumption (e.g. Moor, 2008; Negus, 2002; Smith-Maguire and Matthew 2012, 2014; Wright, 2005; Wright, 2015).

The tastemaking aspect of cultural intermediation is variously explored in different domains, such as fashion (Skov, 2014), advertising (Kelly, 2014) and lifestyle magazines (Lewis, 2014). Cultural intermediaries are assumed to be a part of selecting emerging creativity and making creative production available for group audiences identified as a potential target

(Hutchinson 2015). The concept of a cultural intermediary is also used to study market actors (Callon, 2007; Callon et al., 2007; Muniesa et al., 2007). These scholars use this concept in combination with the insights from the actor-network theory and economic sociology to point out the role human and non-human actors that mediate between the spheres of culture and economy (Webster et al., 2016).

ii. Cultural Intermediaries and the Production of Culture

The cultural intermediary concept is assumed to complement the views on cultural production (Negus, 2012; Smith-Maguire and Matthew, 2014), which focuses on the implications of the processes and practices on the production side of cultural content. These concerns are generally associated with the theoretical approach known as the 'production of culture' (Peterson, 1976). This perspective invites attention to the significance of people who have worked in supporting or gatekeeping roles within media and culture industries. It suggests that workers in specific positions within the organizational structures of media and culture industries undertake formal 'selecting' roles. The embodied expertise of these decision-makers becomes the key arbiter in processes pertaining to the selection of what gets 'commissioned', 'published', or 'promoted', while the forms of knowledge underpinning these decisions become tacit and taken-for-granted modes of professional conduct. In this regard, people working in these strategic gatekeeping and mediating roles are "more readily cast as tastemakers, whose practices remain relatively hidden" (Wright, 2015: p.125).

As this brief review indicates, the notion of cultural intermediaries is a widely used concept for studying a variegated set of actors and activities (Barna, 2017; Webster et al., 2016), described as co-producers, gatekeepers, brokers, agents, matchmakers and tastemakers (Barna, 2017; Foster et al., 2011; Hracs, 2015). This demonstrates that the exact role they perform within value chains is ambiguous (Foster et al., 2011). Nevertheless, as both a theoretical analytic category and a descriptive label for particular occupational entities and contextually specific actors (Smith-Maguire and Matthews, 2012), the notion of cultural intermediaries is considered to provide a number of insights, and it points to some important lines of enquiry. In this regard, Negus (2002) argues that understanding the processes mediating between the production and consumption of symbolic goods requires questioning what seems ostensibly obvious: "Who are cultural intermediaries" and "what is their special position in the relations of production and consumption?"

To address this problem of conceptual ambiguity, Smith-Maguire and Matthews (2012)

define the primary work of cultural intermediaries as constructing value by framing how others engage with cultural content, thereby affecting others' orientations towards content identified as legitimate: "The work of cultural intermediaries is not common to all, because of its expert orientation. In the struggle to influence others' perceptions and attachments, cultural intermediaries are differentiated by their explicit claims to professional expertise in taste and value within specific cultural fields" (Smith-Maguire and Matthews, 2012 p. 2). This view suggests considering the role of cultural intermediaries as not neutral parts of the processes that come in between the creators/producers and the audiences/consumers of cultural content (Negus, 2012). It recognises some of the distinct processes they are involved in between the space of production and consumption of cultural content. In the following section, I review the previous insights on these distinct processes.

i. Curation and Gatekeeping

Curation is considered to be an important subfield of intermediation (Balzer, 2014; Hendricks, 2015). As a concept, it has a wide range of definitions, but it is commonly defined around activities like sorting, evaluating and recommending (Jansson and Hracs, 2018). The term originally used to describe the work performed to organize art collections. But later it was extended to include the selecting, assessing, exhibiting and framing things other than art (Balzer, 2014), such as music (Hendricks, 2015; Jansson and Hracs, 2018), fashion products (Leslie et al., 2015), and so on.

In this vein, Barna (2017) conceptualises curatorship as an increasingly 'professionalised tastemaking and promoting function', which involves 'distinction' and 'representation' and 'exerting control' over choices. Taking a critical stance to online curatorial practices and drawing on the similarities in definitions between *tastemaking* and *curation,* she asserts that curation has become a neutralised marketing term for *tastemaking* and *gatekeeping*, which she defines as "functions that imply the occupation of key strategic positions within a power structure." In this regard, she calls for studies on curatorial practices that take place online, where she argues that "a multitude of various kinds of taste-makers, new and old, seem to coexist in key positions."

From this perspective, curators are also considered to be one of the most important types of gatekeepers. Indeed, the gatekeeping concept is commonly used to understand determinants of how audiences *encounter* cultural content, and, in particular, "what affects the content that

audiences get" (Thorson and Wells, 2015: p.31). Gatekeepers often engage with various types of curatorial work. They are considered to be active selectors and shapers of content, who are not only involved in *filtering* but also *searching out*, and *remixing*, *framing* and/or *reframing* cultural content (Jenkins, Ford and Green, 2013). In this regard, gallery owners, producers in television, editors in publishing houses, or decision-makers in recording companies, and so on are all considered to be strategically significant actors who make decisions on how a particular cultural content becomes available and appears to audiences for consumption.

Bhaskar (2016a) illustrates this through the work of recording companies, which have long acted as significant gatekeepers. He highlights how the selection mechanism regarding which musicians will be promoted has a meaningful impact on the presentation of artists in ways that reinforce the likelihood of being listened to. In the same vein, Wright (2015) argues that the legendary status of most famous and much-appreciated cultural works owe much of their success to the process of *identification*, *selection* and *promotion* of actors working in gatekeeping positions. These processes are considered as being closely related to the curatorial aspects of gatekeeping, which take place before cultural content become available for an audience to choose from.

Moreover, there are other kinds of curatorial work that take place once cultural content is published, i.e. after works are distributed and put into circulation for audiences to choose and consume. This latter form of curation is more about *guiding* audiences and consumers about the quality of available cultural content. In this regard, the curatorial work of gatekeepers in legacy organizations like record labels and publishing companies can be likened to that of DJs, VJs and the like, who work in traditional radio stations or dedicated music channels, e.g. the famous music television station, MTV. These were the prominent cultural intermediaries of their time due to the guiding function of their curatorial work. They had the power to selectively give attention to a few from a range of possible items and direct audiences' attention towards particular cultural content.

In the same vein, critics and reviewers have been influential cultural intermediaries, specifically concerned with guiding audiences in their assessment of cultural content. Although it is difficult to draw definitive boundaries between these concepts as they have many overlapping aspects, for the sake of analytical simplification, I will now discuss the curatorial work assumed by critics and reviewers as separate from, yet complementary to, the work undertaken by gatekeepers.

ii. Critics and Reviewers

The role played by critics and reviewers has been recognized as one of the most powerful industry-based strategies of evaluation used in the assessment of cultural content (Wright, 2015). Critics are described as playing highly crucial roles in apparently autonomous and distinct spaces in the field of cultural production, where "every actor can experience his encounter with the object of his preference as a miracle of *predestination*" (Bourdieu 1984, p. 234; emphasis added). The work of critics is placed within the space mediating between producers and audiences. In that space, evaluations provided by critics and reviewers guide audiences as indicators of 'quality'. Judgements and interpretations, when pronounced by critics and reviewers, considered operating as powerful promotional publicity strategies. For these reasons, they have long been viewed as responsible for shaping cultural tastes (Wright, 2105).

Critics, columnists, and/or editors who write regular reviews in newspapers and/or in various forms of cultural magazines, as well as those figures making TV shows to give recommendations on 'which TV show to watch this week', 'which recently released album to buy and listen to', 'which movie to see this month' or 'whose gig we should go tonight', can all be considered under this category. Accordingly, the mediating role of critics and reviewers in what is referred to as the 'infrastructure of tasting' (Wright, 2015) is assumed to be quite mundane, creating difficulty in apprehension at the moment of choosing to the extent that it looks and feels like we are making discoveries on our own, rather than "selecting between things which are placed before us by, among other social processes, the mechanisms of cultural industries" (Wright, 2015, p.132).

Critics and reviewers owe much of the credibility to their expertise in assessing cultural content and guiding audience based on their demonstrated ability to read audiences' tastes. Relatedly, they are assumed to make forecasts about audience reactions and their future preferences. The legitimacy of the views of such cultural intermediaries hinges on the knowledge claims they are able to make (Smith-Maguire and Matthews, 2014). According to Wright (2015, p.21), making meaningful claims about tastes requires "systematic identification, collation, and translation of individual experiences into data" (Wright, 2015, p.21). But this is not a neutral process.

Producing knowledge about taste is considered to affect the epistemological basis of the

claims made upon this knowledge. In this regard, as an extension of the tastemaking role of cultural intermediation, Wright (2015) suggests considering the ways in which taste is measured and made known as a fundamental dimension shaping cultural taste, which is closely associated with the tastemaking role of cultural intermediaries. Audience/consumer research has long served as a trusted source helping cultural intermediaries, such as advertising and marketing professionals or the gatekeepers working in media and cultural industries. This long-standing interrelationship between audience research and tastemaking is worth exploring as it is a critical element of how cultural intermediaries establish authority as cultural experts.

iii. Audience Measurement and Audiencemaking

Information about audience tastes used to be measured in media and culture industries through the evidence of watching and/or listening (Napoli, 2011). This was then used to shape the future content made available to audiences for consumption. Gaining insight into tastes of their audiences has always helped producers of content to have an advantage over their competitors. The generation of knowledge on the demographic characteristics of an audience - such as their age, sex, education, and lifestyle - also has another crucial purpose for media companies: constructing audience as a commodity to be sold (Smythe, 1981). For this reason, audiencemaking (Ettema and Whitney, 1994) has been one of the primary sources of revenue in media industries, at the heart of which sits the processes and practices of measuring audiences' tastes.

According to Wright (2015), the measurement of taste has always been involved in shaping taste. This is because investigations on taste "require logistical decisions in their design and analysis, as well as assumptions to be made about the kinds of stuff tastes are and how they can be rendered analysable and for what ends" (Wright, 2015, p.42). *A priori* assumptions about the categories used to identify taste used to involve particular epistemological suppositions about the notion of taste and how it can be observed and measured. *A posteriori* forms of measurement complemented this through instruments such as questionnaires and surveys, which rely on subjective accounts of the respondents, as recalled from past experience. As Wright (2015) highlights, these techniques have long served media industries in making programming decisions and promoting particular cultural content. The survey method was gradually replaced by other techniques that help media industries gain insight about the tastes of audiences, not only for content production, but also for constructing

audiences as commodities to be sold to advertisers. Information about the habits and preferences of audiences was used to manage demands for and consumption of cultural content, while rendering audiences quantifiable and commodifiable products to be sold to advertisers to provide them with precise targets (Bermejo, 2007; Smythe 1981; Ettema and Whitney, 1994).

Knowing audiences has always been an issue for the media and entertainment industry, and audience has always been constructed in different contexts for different purposes (Ang, 1991). Knowing more about the tastes of audiences has helped cultural intermediaries working at media companies to fine-tune their programming decisions (Wright, 2015). Using audience data to reduce uncertainty and increase the predictability of audience behavior (Napoli, 2011) was an important element of mass media logic (Van Dijck and Poell, 2013). Statistical methods and quantification have long served as the trusted tools of audience measurement (Ang, 1991) from which to make predictions about the behaviours of an audience. These techniques are assumed to reduce the differences in diffusion of socio-cultural contexts and differences in individuals' reception of cultural content. In this way, the differences among individuals watching television, listening to radio, or reading the newspaper can be turned into a homogeneous mass of measurable objects, which can then be erased and distinguished by a defined set of criteria (Passoth, Sutter and Wehner, 2014). As a result, what is known as audience, the assumed role of a viewer, listener, or reader, and how they become distinguishable from each other have all been considered as linked to the processes of measuring the audience (ibid.)

There is extensive literature that critically engages with audience measurement practices. In his seminal work on media audiences, Smythe (1981) famously reveals that audiences are actually at work when they are watching television. In the course of watching advertisements, audiences are generating data about their preferences to be sold to advertisers. In the same vein, Meehan's (1984) influential work has been particularly significant and pioneering in reframing Smythe's (1981) audience commodity thesis, turning attention to the work of *measurement technologies*. Meehan (1984) argues that audiences are produced as 'knowable entities' through the ratings used to measure them. What audience measurement companies like AC Nielsen create is only a partial version of what constitutes the actual audience because the measurement companies are trying to produce audience information in a way that would be appealing to the companies relying on that information. Meehan's (1984) work

reveals that, rather than being a true representation of audience members, audience commodity is composed of 'figments of measurement' that are used to convince advertisers. In making a significant contribution to literature by updating Smythe's (1981) formulation, Meehan's (1984) work highlights how audiences are, in fact, constructed, and the measurement technologies are part of this construction. In this regard, techniques and technologies involved in the measurement of audiences are argued to become constituents of the final audience product. Despite often being assumed to be constituted of actual people watching or listening to cultural content, the term 'audience', according to Ettema and Whitney (1994), refers to a specific product category that is constructed by audience measurement arrangements.

As the term 'audience' refers to the 'commodity' produced by media companies, the notion of 'audiencemaking' describes the very 'process' in which audiences are constructed as commodities. In this regard, Bermejo (2009) asserts that measurement technologies are quantitative contractions of audiences that manufacture *a particular kind* of audience. To illustrate, Ettema and Whitney (1994) demonstrate how, rather than reporting an actual viewing audience, rating firms and measurement technologies provide 'institutionally effective' audiences. The audience becomes institutionally effective when it can be efficiently integrated into the economic transactions occurring in media industries. In this regard, some scholars critically engage with the implications of measurement practices in terms of how knowledge produced about audiences can variously feedback to shape their choices (Wright, 2015).

Taken all together, previous scholarly works emphasize that the process of measurement does not generate a neutral description or the translation of a reality about audience members 'out there', but rather these measures are, in fact, constitutive of what they purport to describe; that is, the audience taste. Audience members are classified under certain categories, such as genre types (Lena, 2012), based on the assumptions embedded in measurements and the ways in which they take place. These arrangements used to be built as top-down taxonomic structures that are generated based on what is considered expert knowledge (Gillespie, 2014); that is, they used to be in the preservation of specialists, such as critics, academics, and other professional actors that work in the cultural industries. Thus produced, audience categories have long affected how meaningful claims are made about audiences, thereby influencing what content audiences get. In turn, audience reactions to what they listen to, watch, or see

are measured again to reiterate the same process and thus cement the kind of knowledge produced on audience taste (Wright, 2015). Such an understanding is not limited to audience measurement literature. Scholars who are concerned with understanding the sociological aspects of the processes of quantification, categorization, and measurement have also argued that these processes entail significant implications (i.e. Hacking 1986).

2.2.2. Cultural Intermediation as a Process

Previous sections highlighted how the concept of a traditional intermediary has been variously used to investigate different practices in different contexts. It was also discussed how various techniques and technologies have been involved in the processes that accompanied the work of cultural intermediaries, such as the measurement tools that are used to make knowledge claims about cultural tastes. This indicates that it is difficult to draw the boundaries between different types of work that can all be subsumed under the mediating roles undertaken by cultural intermediaries. Drawing on the aforementioned debates about the imprecision of the concept of a cultural intermediary, Powers (2015 p.122) suggests shifting the analytic gaze from 'intermediaries' to the 'process of intermediation'. Such a view contrasts with the literature concerned with cultural intermediaries in which it is often considered as someone's job. Powers' (2015) suggestion is worth paying attention to since the majority of studies on cultural intermediaries focuses on the role of humans and neglects the intermediary roles of technologies, objects, and devices. An exception to this is Moor (2012), who establishes ratings and related measurement techniques as important mediators between production and consumption.

From this perspective, I conceptualise cultural intermediation in this study as a process mediating between the production and consumption of cultural content. It is composed of the curation, assessment, and guidance work that has been performed by cultural intermediaries such as the gatekeepers, critics, and reviewers. It also involves the audience measurement techniques used to make audience tastes knowable. In this regard, based on a synthesis of the literature, *curation*, *guidance*, and *measurement* are conceptualised in this study as the key instances of the cultural intermediation process.

2.2.3. The Transformation of Cultural Intermediation

Digital technologies have become integral to individuals' everyday routines. As such, they are now also centrally embedded in the everyday practices of 'choosing' and 'tasting'

(Wright, 2015). This has altered the existing modes of production, circulation, and consumption of culture, which facilitated a new kind of cultural abundance, and new ways have been generated to cope with it. Thus traditional cultural intermediaries, such as curators and gatekeepers, began to be complemented, and at points replaced, by less formal actors like individual media consumers, amateurs, social media users, and entrepreneurs (Ashton and Couzins, 2015; Barna 2017; Bruns, 2011). These actors engage with the practices of gatekeeping, making recommendations and providing guidance with their curations in the digital space.

These forms of guidance are assumed to be complemented by 'algorithmically' organized modes of recommendation and circulation (Wright, 2015), which Gillespie (2014) calls 'algorithmic logic'. This logic is considered as replacing the 'editorial logic' underpinning the choices of traditional cultural intermediaries, as production and circulation of cultural content organized around the recommendation and ranking systems of online platforms. This indicates that there is also an increased recognition regarding the role of sociotechnical actors in the organisation of culture, such as algorithms, music recommender systems, and social media platforms (Lange, 2016; Webster et al., 2016; Thorson and Wells, 2015). These sociotechnical entities are considered to be increasingly taking over the roles that used to be performed by the more traditional cultural intermediaries.

To illustrate, personalization technologies and recommendation systems have gained increasing recognition in terms of how they reconfigure the work that has traditionally been carried out by cultural intermediaries. They do so by attending to the processes of curation, guidance, and measurement, and thus they increasingly participate in shaping tastes (Beer, 2013; Morris, 2015a; Webster et al., 2016; Wright, 2015). These scholarly accounts point out that sociotechnical systems have significantly transformed the process of cultural intermediation, and by extension they are becoming increasingly involved in tastemaking. These works have generated valuable insight and awareness regarding the sociotechnical aspects of the current transformation of cultural intermediation.

Scholars who acknowledge such implications highlight the role of algorithms (Beer, 2009; Bruns, 2011; Lange, 2016) and/or the recommender systems (Beer, 2013; Morris, 2015a; Webster et al., 2016) of which they are a significant component. Algorithms and recommender systems are certainly important, but they do not operate in a standalone fashion. It is important to recognise that they are embedded within widely distributed and

standardized digital data sources (Kallinikos and Tempini, 2011) that are increasingly involved in mediating the social and cultural relationships (Kallinikos, 2006). Social media is a prominent example of this. In this regard, the next section reviews the literature on social media and the current views on how social media are involved in the transformation of cultural intermediation.

2.3. Literature on Social Media

Extensive research on social media stems from the significant role they play in contemporary daily life. Well beyond their original intent of connecting individuals, social media have transformed the way people, organizations, brands, businesses and governments communicate. By rendering social interaction and cultural consumption incredibly fast, easy, and convenient, they have become the new means of self-expression and social interaction. Social media has enabled ordinary individuals to become content creators, and enabled both individuals and organizations to reach out to new audiences that they never would have been able to communicate with otherwise. In doing so, they have also altered the ways in which cultural content is created, distributed, discovered, and consumed. Thus social media have significantly reconditioned how culture is produced, circulated, and experienced (Beer, 2013).

Social media generate a tremendous amount of personal data in mediating, as well as to mediate, the aforementioned processes. This personal data – digital data created *by* and *about* people *on* and *by* social media – is considered to have generated a new wave of opportunity for economic and social value creation. Mining and analysing the data is increasingly used in the service of understanding behaviour and predicting future activity at an individual, group, and even global level. Data is the fuel of the future and is giving rise to a new economy (van Dijck, 2014; van Dijck and Poell, 2013). A considerable portion of this data is generated by and on social media as a by-product of user platform participation. Due to its rising social and economic importance, and because the transformations enabled by social media have implications that are so far-reaching, there is a growing and diversified body of scholarly work aiming to build an understanding of social media. In the next sections, I will review the key insights from the literature on social media and illustrate how the conceptualisation of social media has evolved over time from networking sites to platforms infrastructuring daily life.

2.3.1. Social Networking Sites

Early views on social media in media studies and sociology define them as 'networking sites' (Boyd and Ellison 2007; Beer 2008), which are "web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system" (Boyd and Ellison, 2007, p. 211). This approach analyses social media mostly as user-enabled social spaces constituted of static and closed systems. These accounts consider notions of identity, self-presentation, community, and culture as the central themes of inquiry (Papacharissi, 2010; Baym 2010; Kietzmann et al., 2011). Over time, this 'networking sites' approach evolved to delineations describing social media as neutral facilitators of social interaction, fostering communication in ways that allow the creation and exchange of 'user-generated content', which is defined as all sorts of digital material generated/shared by users (Kaplan and Haenlein, 2010; Kietzmann et al., 2011). They are thus conceptualized as 'archives' of the everyday (Beer, 2013; Zhao, Salehi et al., 2013) built around 'prosumption' activities proliferating in the cultural sphere. This result in the creation of new forms of social data generated as a by-product of online social and cultural interaction (Beer and Burrows, 2013).

This notion of 'prosumption' is widely adopted by social scientists to explain the simultaneity of the production and consumption of content (Beer, 2009; Beer and Burrows 2010; Fuchs, 2014; Ritzer et al. 2012; Zajc, 2015; Fergie et al., 2016; Chen, 2018). These accounts point out how, with the rise of social media, ordinary people have been increasingly participating in the formation of media content, which results in vast amounts of data generated about individuals. In these accounts, cultural data generated on social media are often regarded as naturally occurring by-products of online interaction and communication (Beer and Burrows, 2013; Beer and Taylor, 2013; Boyd and Crawford, 2012; McAfee et al., 2012), while social media are considered as semantic 'taste fabrics' (Liu et al., 2008), affording novel ways for representing and reasoning about web users' deeper patterns of taste. In such views, the data that social media produce are significantly taken for granted as straightforwardly produced true mappings of social and cultural dispositions and reliable indicators of aspects of an individual's identity.

These initial conceptions of social media as networking sites have gradually evolved and been complemented by more critical approaches that engage documenting the structural and

technological complexity of social media in relation to the economics and politics of data. Thus earlier approaches have since been problematized on the grounds that they tend to reduce social media to what is happening in the front-end, on the interface level (Couldry and Kallinikos, 2017). Critiques of such views instead argue that social media are artificially created technological environments that are designed with specific business and commercial objectives (van Dijck, 2013; Langlois and Elmer, 2013; Gillespie, 2010). And, contrary to views that see data generated on social media as fine grained-data *about* the social (i.e. see Beer and Taylor, 2013), 'social data' is defined as data produced *by* social media *through* the engineering of user-platform participation (Alaimo and Kallinkos, 2016; 2017).

Similar concerns are raised also in relation to the political economy of online platforms. In this vein, social media are conceptualised as data firms or data-based organisations that engage in the harvesting of personal data generated by the datafication of mundane activities (van Dijck & Poell, 2013; van Dijck, 2014). According to the definition of Mayer-Schoenberger and Cukier (2013), 'datafication' refers to the transformation of social action into online quantified data, presented as 'raw material' that can be processed for analysis to predict future behaviour. This notion of datafication of mundane human activity, according to van Dijck and Poell (2013) constitutes the distinct logic distinguishing social media from the earlier forms of media, which has resulted in an industry that builds on the value of data and metadata (van Dijck, 2014).

More and more aspects of social life are now coded into data through the quantification of online behaviours and interactions while this data made accessible to third parties for repurposing. This has led to the proliferation of online platforms that offer social networking and content-based cultural services for free in return for the personal behavioural data voluntarily provided by users. These developments, according to van Dijck (2014), have rendered personal information and behavioural data the new currency to be paid in return for free services provided by online platforms like social media. Accordingly, the behavioural data produced on social media is considered to turn into a kind of invisible asset, processed mostly as separate from its original context and outside of people's awareness in order to construct targetable markets for marketing and advertising (van Dijck and Poell, 2013; van Dijck, 2014).

From this perspective, the data accumulated on a massive scale as by-product of online individual and group behaviours produce a new form 'social data' *about* the social, which

allow micro and macro level analysis, and, thus, have the potential to reveal 'hidden dimensions' and make it possible to 'see the world differently (Beer and Taylor, 2013). Thus social media data have become increasingly acted upon sources that are assumed to reveal previously unknown patterns hidden in the data, which, according to Kallinikos (2013), are assumed 'to lift the veil of reality'. Scholars who challenge such neutral depictions of the nature of social media data, on the grounds of problematic ontological and epistemological assumptions underpinning their generation and use (Boyd & Crawford, 2012; Gitelman, 2013; van Dijck, 2013, 2014; Kitchin, 2014b; Couldry and Kallinikos, 2017), have paved the way for the more recent conceptualisation of social media as platforms.

2.3.2. Platforms Infrastructuring the Everyday

The concept of platform is used to understand a diverse range of issues regarding social media as platform businesses (Gillespie, 2010; Langlois and Elmer, 2013; van Dijck, 2013; van Dijck & Poell, 2013; Helmond, 2015; Plantin et al., 2016). For instance, Gillespie (2010) argues that the term 'platform' brings together and aligns different discourses about businesses built on platform-based business models. He argues that this notion helps these businesses to sustain different sides (markets) that they serve by using different kinds of rhetoric that appeal to their users and customers. Beyond the discursive aspects, there are also studies that conceptualise social media platforms as 'socio-technical arrangements' (Burgess et al., 2017; Bucher and Helmond, 2017) affording and constraining activity (e.g. Baym, 2010; Ellison and Vitak, 2015) in ways implicated in shaping online forms of communication and expression.

Social media are argued to create a distinct form of 'connectivity' (van Dijck, 2013) which necessitates considering social media affordances in relation to the larger context within which social media operate. In this regard, van Dijck and Poell (2013) invite attention to the ways in which 'the socio-technical affordance' of social media is involved in connecting user activities and advertisers around platform content. From this perspective, social media platforms are seen as rapidly evolving digital artefacts shaping online interactions in line with their distinct business logic and commercial objectives (van Dijck, 2013), which depend on social media's ability to make users relevant to each other (Gillespie, 2014) through their matchmaking operations (Parker et al., 2016).

As an extension of this perspective, social media platforms are considered to resemble ‘infrastructures’ (Alaimo and Kallinikos, 2017; Helmond, 2015; Plantin et al., 2016; Gillespie, 2010; van Dijck, 2013) that are taken for granted and become invisible yet are significantly reconditioning everyday living. At the centre of the concerns regarding the infrastructuring role of social media lies the barely understood implications of the digital transformation of social and cultural interaction. From this perspective, social media are viewed as a ‘functional infrastructure’ for users’ everyday activities, which are “not simply channelled or facilitated by platforms but programmed in a determined way and with a specific objective” (van Dijck 2013, p.6). These scholars stress the need for studies examining the technological enabling and codification of communication acts on social media, a process that van Dijck (2013) refers to as 'plat- formed sociality'. Langlois and Elmer (2013) refer to this process as ‘platforming’ in order to advocate moving attention away from what is happening on the interface level (what users say and post as part of identity construction and presentation of self, etc.) to how these are processed and rendered computable through the production, storage and processing of data, which results in social media’s infrastrucuring of the ‘remaking of everyday’ (Alaimo and Kallinikos, 2017).

To conclude this section, it is important to note that, despite conceptualising social media differently, reviewed perspectives indicate that social media facilitate the interactions between individuals and cultural content quite differently than other media did in the past. This raises questions regarding the implications that this has for the creation and consumption of cultural content, as well as for the processes mediating between these two ends. These questions are far from trivial, given the extent to which social media have penetrated into the everyday living of individuals and the significance they hold for the wider digital economy. The next section will elaborate on this by reviewing the literature concerned with how social media are altering cultural intermediation.

2.4. Cultural Intermediation and Social Media

In this section, I review existing works that have examined how social media shape curation, guidance, and measurement, the processes conceptualized in this study as the prominent instances of the cultural intermediation.

2.4.1. Curation in Social Media

According to media scholars, the fundamental action in the digital media environment is curation - that is, the 'production', 'selection', 'filtering', 'annotation', and 'framing' of content (Thorson and Wells, 2015). Curation is seen as an activity that has always been an integral process in the dissemination and consumption of cultural content. Hence it is an important instance of the cultural intermediation process (Jansson and Hracs, 2018). Digital affordances have enabled anyone to publish content online and the problem 'cultural scarcity' of the past is replaced by the problem of 'cultural abundance' (Bhaskar, 2016a; Wright, 2015). The age of cultural abundance has created a new obstacle in the realm of culture, *discoverability*. Consumers become overwhelmed by the amount of choice and information available to them. As a result, audiences have started to be in need of help to make sense of the cultural content circulating in the digital space. Thus intermediaries have become more prominent than ever. In this regard, scholars observe that curating cultural content has become more important and more valuable than creating it (Ashton and Couzins, 2015; Balzer, 2014; Jansson and Hracs, 2018; Bashkar, 2016a, 2016b).

Curators help audiences by identifying what is valuable and excluding what is not, thereby cutting down complexity, saving time, freeing up cognitive resources, finding quality and ultimately overcoming 'information overload' (Bhaskar, 2016a). Unlike in the mass media era, wherein access to content is largely framed and tightly controlled by powerful media actors, curation in the digital context is seen as undertaken by a set of new actors (Thorson and Wells). End-users, constituted of ordinary individuals, have become 'half creator' and 'half curator' of the cultural content circulating in digital contexts. Consequently, the organization of music has been considered to shift away from the straightforwardly top-down industrialised organization of what used to be known as the 'broadcast model' to a 'consumer-curated model' (Bhaskar, 2016a): "*The power to decide' who listens to 'what and when' has shifted from broadcasters to audience. Albums and radio are broadcast; playlists are consumer-curated.*"

Social media is considered to play a significant role in this shift. Numerous kinds of social media platforms built around social curation and discovery are launched with the mission of helping users find new content. Social media's influence on curation practices has also been investigated, primarily to shed light on how established processes of professional curation are being altered in the digital context. For instance, Bhaskar (2016a) studied in the domain of

music discovery to investigate this. He argues that the evolution of music discovery in this trajectory has resulted in the blending of *consumption* with *curatorial* practices (Bhaskar, 2016a): Once there were only recorded albums providing listeners with a *set order* and *selection*. Then listeners themselves could prepare mix-tapes for their loved ones. Nowadays, users act like DJs, creating playlists by curating songs in a particular order according to their moods, genres etc., and sharing those playlists online. Today they have the opportunity to *showcase* their music taste to a global audience in music streaming platforms like Spotify, which increasingly give more space to their users to create and share *personally curated playlists* that *showcase their music tastes*.

iv. ***'Social Curators', 'Gatewatchers' and 'Community-driven Curation'***

Closely related to this trajectory of discovery in the digital context is the term 'social curator' (Villi, 2012). This term is used to describe ordinary individuals who aggregate, share, rank, compare, critique, and distribute content on social media platforms. At the core of social curation lies something users do primarily *for others*. As Villi (2012) emphasises, this is not just about sharing interesting and useful links; rather it is 'selecting', 'collecting', 'organizing', and/or 'archiving' handpicked content. It involves the judgment of cultural content with the aim of showing as a means of expression on social media. Such user-curation is regarded as the *showcasing* of *users' tastes*. Sharing curated content as a form of showcase is also considered to function as a form of *guidance* and *gatekeeping* (Villi, 2012): online sharing depends on providing a link for the recommended content, which entails the provision of access to it.

In this regard, a 'social curator' is defined as someone who monitors, finds, selects, and organizes the most relevant content for a specific community of audiences. Because social curators give their audiences access to the content they recommend, they function as gatekeepers. Stanoevska-Slabeva, Sacco and Giardina (2012) illustrate this point by referring to these social media curation activities as 'gatewatching'. They describe this as crafting digital narratives out of online and social media content. As social media platforms proliferate, curators on social media become the 'gatewatchers' who undertake previous forms of gatekeeping roles. A more specific example comes from Bruns (2015, 2018), whose work highlights how news curation constitutes one key use of social media at present. News curation on social media, as she argues, is a new practice of communal news curation that builds on individual participants' gatewatching efforts. These efforts are seen as a form of

(citizen) journalism, and news curation driven primarily by information sharing activities is assumed to foreground the input rather than output aspects of that journalistic process. In Brun's (2015) words: "It is at its core more an exercise of research and compilation than one of interpretation and publication" (p.12).

Parker et al. (2016) highlight another aspect of such communal practices enabled by online platforms in general, including social media. They argue that platforms have reconfigured quality control through the curatorial work of a community of users and the reactions the curated content gets from other users. They call this 'community-driven curation': Platforms create the infrastructure that enables content producers to create and circulate content - may it be a photo on Instagram, a video on YouTube, or a status update on Facebook. In doing so, they also foster a culture of quality control in the form of quantified and aggregated user reactions materialised through the visible platform metrics (i.e. number of 'likes' or comments a content gets on Facebook, the thumbs-ups a video gets on Youtube, or rankings of various kinds).

v. Tastemakers are all Around - They Just Need To Be Connected!

As these accounts illustrate, the dominant tendency in theorising the transformation of curation and gatekeeping practices is to put emphasis on the emergence of a new type of cultural intermediaries constituted of ordinary individuals using social media. The main assumption is that, if the work of curation involves the framing and contextualization of cultural content for a specific group of audiences, and gatekeeping relates to what is selectively made visible to the gaze of audiences, then all social media users work as contemporary forms of curators and gatekeepers. As Mueller (2012) puts it: "*In social media, we are all gatekeepers and curators*." Each status update, each post, each pin and tweet is a 'digital object' that users create, curate, or share. "*As we see things around the web, we make a determination as to which of these objects are worth sharing. We 'like', comment, share, tweet, pin, and so on. Each time we do that, whether we provide context or not, we are endorsing that object*."

The view is that, when bringing together a collection of cultural content in a specific order (e.g. user playlists in YouTube, user boards on Pinterest, etc.) on social media, users provide context and frame cultural content in particular ways, and this thus addresses the aforementioned problem of discoverability resulting from the information overload or content over-abundance in the digital realm. Thus, by handpicking and bringing cultural content

together from different sources to exhibit in a central space, social curators are assumed to play a key role in shaping how others encounter cultural content. In this regard, Jenkins et al. (2013) assert that 'grassroots intermediaries' are taking over the role of traditional cultural intermediaries. As an extension of this view, ordinary individuals who participate in social curation are seen to operate as contemporary tastemakers who can reach out and guide a global set of audiences using social media. As Morris suggests: "*Tastemakers are all around - they just need to be connected*" (Morris, 2015b: p.118).

2.4.2. Guidance in Social Media

Guidance is another key instance of the processes mediating between the production and consumption of cultural content. Therefore, within the context of this study, guidance is established as an instance of cultural intermediation. Previously, I discussed how previous forms of guidance entailed the assessment and review of cultural content by critics and reviewers (Wright, 2015). Their views on the quality of cultural content, and their recommendations on which ones are worth seeing, listening to, or reading, have played significant roles in guiding audience tastes in the past. In this section, I will review the literature with respect to its implications on changing forms of guidance that an audience gets on social media.

The most prominent example of how social media alter traditional forms of guidance is personalized recommendations. This process is assumed to constitute a novel form of guidance enabled by social media platforms. Ordinary daily experiences are now continuously recorded on social media to construct 'taste profiles' (Alaimo and Kallinikos, 2016; 2017). Thus the platform decides which content to be shown to which user based on the calculated relevance between the users and various content circulating on the platform, and it produces personalized recommendations *matching* each user's *taste.* This processes is termed as 'algorithmic curation' (Bhaskar, 2016a) underpinned by 'algorithmic logic' (Gillespie, 2014), as opposed to the traditional 'editorial logic' built on expert choices.

Profiling users' tastes, predicting their future orientations, and providing them with corresponding recommendations have become central to social media (Alaimo and Kallinikos, 2016; 2017). Thanks to the work of their recommendations systems, these platforms have increasingly been trusted sources for the access and discovery of cultural content and are delegated a great deal of autonomy for determining what is relevant to users' taste. For these reasons, social media's recommendation systems have gained increasing

recognition in terms of how they reconfigure the work that has traditionally been carried out by cultural intermediaries, and hence, how they attend to shaping tastes (e.g. Beer, 2013; Morris, 2015a; Bhaskar, 2016a; Vanderbilt, 2016; Webster et al., 2016). Some scholars attribute the tastemaking aspect of recommendation systems to a new hybrid constituted of human and non-human or machine choices (Bhaskar, 2016a; Webster et al., 2016), while others attribute this to the emergence of new user practices, or to the work of algorithms.

For instance a recent work on the role of social media's recommendation systems in taste formation studied it around emerging user practices (Karakayali, Kostem, & Galip, 2018). Findings of this study indicate that social media users draw on personalised recommendations provided by the platform to examine and transform aspects of their identity, such as their 'music taste'" (ibid.). Other studies have focused on the recommendation systems and their algorithms to stress how they transform 'cultural intermediation' and are involved in tastemaking. To illustrate, Beer (2013 p.94) argues that they shape tastes: "recommendation algorithms draw the attention towards cultural products and thus exercise power to shape cultural encounters that then feed into taste." Likewise, Morris (2015a p.450) highlights how algorithms and recommendation engines are operating as a cultural intermediary: "If cultural intermediaries such as critics, talents scouts and the like were once responsible for the 'presentation and representation' of culture, or for the construction of value and framing of encounters with cultural goods, it is hard to deny that recommendation services, and the algorithms that constitute them, are increasingly part of the intermediation process.*"*

According to Morris (2015a), algorithms mediate culture by processing data about listening behaviour and musical content, and they "*shape taste* and *derive legitimacy* in a different manner than cultural intermediaries" (Morris, 2015a p. 456 emphasis added). The ways in which this legitimacy is gained are, according to Wright (2015), significantly linked to the algorithms' involvement in *how* taste is *measured* and *made known* in digital environments. Attributing the aforementioned processes of constructing taste profiles (Alaimo and Kallinikos, 2016; 2017) to the work of algorithms, Wright (2015) describes algorithms' participation in taste formation as the following: "*Algorithms aim to render our tastes and preferences visible to gazes outside the immediate moment of tasting principally to the advertisers and marketers for other products for which we are potential audience*" (Wright, 2015, p. 158).

Building on such views, Bhaskar (2016a) contends that, despite increased automation, the necessity of human choice remains to be the central guiding element in social media: "*The more we have, the more we rely on algorithms and automated recommendation systems. Hence the unstoppable march of algorithmic recommendations, machine learning, artificial intelligence, and big data into the cultural sphere… Yet this isn't the end of the story… Far from disappearing, human curation and sensibilities have a new value in the age of algorithms. Yes, the more we have the more we need automation. But we also increasingly want informed and idiosyncratic selections. Humans are back*." This is considered to be true especially in the domain of culture. According to Bhaskar (2016b), the cultural sphere is distinct from other sectors taken over by technological disruption in the sense that the 'human touch' is an irreplaceable element for its operations: "*We want to be surprised. We want expertise, distinctive aesthetic judgments, clear expenditure of time and effort. We relish the messy reality of another's taste and a trusted personal connection. We don't just want correlations – we want a why, a narrative, which machines can't provide. Even if we define curation as selecting and arranging, this won't be left solely to algorithms*."

As these accounts indicate, the predominant approach in explaining the role of social media in the transformation of cultural intermediation is to attribute the novelty to the involvement of ordinary individuals and the automated algorithms of the recommenders systems in this process. These aspects are closely linked with *how* taste is *measured* and *made known*, a significant aspect of tastemaking (Wright 2015). The next section reviews the literature on social media with respect to their implication on measurement of taste.

2.4.3. Measurement in Social Media

In digital contexts, audiences can create and curate, and they also have increased control and choice over 'when', 'where', and 'how' to consume media and cultural content. This has given rise to the emergence of new audience measurement techniques that allow real-time tracking and analysis of audience behaviour, which is assumed to reveal previously unknown aspects of audiences' media consumption and corresponding cultural tastes. This, according to Napoli (2011), indicates the changing dynamics of how audiences consume media and the ways in which their behaviours are monitored and measured in digital environments, such as social media. In this regard, social media are considered to pave the way for a fundamental transformation in how audiences are made knowable entities. This transformation started with

the increased involvement of computational technologies in audience measurement and audiencemaking and taken forward by social media.

Computation has radically transformed traditional audiencemaking processes (Aaltonen, 2011) by shifting value creation from "obtaining valid and reliable samples of people's media consumption to analysing the audience from the extant data" (Aaltonen and Tempini, 2014, p. 106). Similarly, social media have altered various aspects of audience measurement, such as measuring 'engagement'. The most central form of audience measurement has been counting audience size through sales or exposure (Ang, 1991). But audience members were observed to be more engaged than what these measures captured. Yet, the technology at hand was not able to measure dimensions of engagement in 'institutionally effective' (Ettema and Whitney, 1994) ways. Before the interactive possibilities of Web 2.0, measurement of user engagement was limited to counts of hits and time spent on webpages (Gerlitz and Helmond, 2013). Digital convergence has led media products to be offered as services[1] through online platforms, while interactive features of these platforms have rendered audience engagement increasingly visible and traceable.

Over time, this aspect has paved the way for seeing social media as a rich source of information about users due to its perceived ability to measure audience engagement with content. Thus the daily preferences, habits and affects of audiences have started to be measured in ways that extend far beyond the reach of previous forms of consumer and audience research conducted in the past. The most distinctive aspect of the innovation brought to audience measurement by social media is the way in which individuals' tastes are observed, captured, and acted upon (Alaimo and Kallinikos, 2017). Social media generate new types of data about the behaviour of audiences. It is different from the data generated about audiences in the past using demographic classifications and the social categories they signify, which have long served as proxies for consumption patterns. These traditional classificatory methods are rendered less relevant in the context of online spaces such as social media, wherein people actively indicate their interests (Aaltonen, 2015). As long as the audience is engaged, the concern is no longer knowing '*who she is'*, but, rather, knowing more about '*what she does'*.

[1] This phenomenon is often referred to as a shift from a product-dominant logic towards a service-dominant one (Vargo and Lusch, 2004).

This has shifted attention towards measuring different dimensions of engagement. Measuring the depth rather than the breadth of engagement has become more important. So social media platforms have started to particularly look for 'active', 'engaged', and 'influential' audiences (Napoli, 2011). Interactive features of social media have also altered the nature of engagement, as they make it possible to gather feedback from the audience. As a result, social media has emerged as the primary means through which audience engagement with content can be translated into performance metrics used to assess and valuate audiences (Napoli, 2014). In this regard, aforementioned game-like social media features - including active audience choices to 'like', 'follow' or 'share' particular cultural content - function as proxies for measuring 'size', 'engagement', and 'affect' (Gerlitz and Helmond, 2013). Such social media metrics have become heavily relied upon sources of audience engagement (Baym, 2013), while also being increasingly taken up as strong indicators of users' dispositions and preferences. Higher counts on these metrics are assumed to represent larger and more engaged audiences, which, in turn, are expected to stimulate further engagement from other users (ibid.).

These accounts suggest that social media have transformed audience measurement and audiencemaking in significant ways. This is closely related to how audience tastes are measured and made known to be used in operations of social media. This may include repurposing data to create targeted market opportunities (van Dijck, 2014; van Dijck and Poell, 2013) for advertisers by rendering particular advertisement content relevant to particular users by means of constructing and delivering personalised recommendations. The same personalisation mechanism underpins the ways in which aforementioned consumers' curations are delivered to users to guide their preferences. Some scholars explain this mechanism using the terms such as 'algorithmic logic' (Gillespie, 2014) or 'algorithmic curation' (Bhaskar 2016a), and they stress the role of algorithms (see also Beer, 2013; Wright, 2015) at the intersection of the notions of cultural intermediation, audiencemaking, personalised recommendations, guidance, tastemaking, and so on.

Algorithms are certainly important, but they do not operate in a standalone fashion. In order for an algorithm to provide personalized results, users' taste must first be rendered data and readied for the algorithm, as is also emphasised by Gillespie (2014). The current state of literature on social media in relation to audience measurement and audiencemaking does not

provide a detailed explanation of how exactly audience tastes are captured and turned into data.

2.5. Conclusion

This chapter provided a review of the literature on cultural intermediation and social media, highlighting how their conceptualisations evolved over time. Against this backdrop, the chapter reviewed the current views on how social media are involved in the transformation of the cultural intermediation process by having a closer look at curation, guidance and measurement in social media as its key instances. This review highlights that the mediators of the past (i.e. gatekeepers, curators, and reviewers), who are considered to have 'collectively dictated' what content should be consumed, have now given way to a complex mixture of *algorithms* and *user-curations* (Bhaskar 2016a).

These accounts illustrate that the predominant approach to understand how social media transform cultural intermediation is attributing the novelty to the work of of users or the algorithms. Algorithms are certainly important, but they do not operate in a standalone fashion. In order for an algorithm to provide personalized results, users' taste must first be rendered data and readied for the algorithm (Gillespie, 2014). Therefore, that abstract concept called 'taste' first needs to be decomposed into its elementary units in ways that enable it to be captured and rendered into computationally processable data. Nevertheless, overemphasising the role of algorithms and users result in neglecting the role of data in this transformation.

This is an important gap given that algorithms act on the data-rendered taste to guide and assist audiences by delivering them the aforementioned user-curations in personalized ways. This gap cap should be addressed by laying bare how audience taste is captured and rendered into computational data to be measured, known, and acted upon in order to provide audiences with personalized recommendations matching their taste. Previously reviewed social media literature indicates that the sociotechnical design aspects of social media significantly shape frontend users practices, such as the aforementioned user-curations, in ways that help the construction of personalized recommendations at the backend (Alaimo, 2014; Alaimo and Kallinikos, 2016; 2017; Bucher, 2012; van Dijck, 2013). These works focused attention on laying bare the inner workings of social media to open up a critical discussion regarding its possible implications for social and cultural interactions. Nevertheless, there is still a growing

need for further empirical research (Couldry and Kallinikos, 2017) to advance our critical engagement with the implications of these developments. In this regard, this study aims to go beyond the current explanations emphasizing the role of users and algorithms and have a closer look at the sociotechnical configuration of social media to understand their role in the transformation of cultural intermediation.

3. Conceptual Framework

3.1. Introduction

This chapter presents the conceptual framework that serves as the basis for the case study. It defines the concepts that will be central for the analysis, discusses how they have been used in previous research, and proposes relationships between them. The proposed framework serves as a less developed form of theory that will be tested and developed in the analysis of the case study findings. From the conceptual framework, more specific research questions are derived.

The chapter is structured as follows: I first provide the background for reviewing the literature on platforms in light of the notion of platformization, which symbolises the transformation of linear flows into networked arrangements. Following this, I review the literature on platforms. First, I draw on management and economic studies and characterise digital platforms as multisided markets reliant on network effects. They are conceptualised as 'matchmakers' on the grounds of the distinct matchmaking logic that allows them to facilitate interactions among external participants. These views are complemented by insights from another stream of research focused on the distinct sociotechnical configuration of digital platforms, and in particular social media as their pioneering instance. This strand of research is concerned with understanding how social media shape social and cultural interactions on the frontend in ways that generate data for their backend data operations.

Drawing on these perspectives, I characterise social media as a prominent instances of digital platforms built on a particular economic logic and sociotechnical design rationale. In this vein, I conceptualise social media as distinct kind of platform that is particularly reliant on continuous user activity on the frontend to generate data influx at the backend for matchmaking operations. Against this backdrop, I elaborate on the distinct relationship between the frontend and backend of social media platforms and how it enables the construction of users' taste profiles to fuel matchmaking operations. Finally I move on to discuss how social media afford 'plat-forming' everyday living. From this perspective, I discuss the platformization of cultural intermediation. Synthesising previous work, I conceptualise cultural intermediation as a process with curation, guidance and measurement processes as its instances. It is argued that, as an extension of platformization, this process inherits the tastemaking role of traditional cultural intermediaries, defined as professional tastemakers. From this vein, the platformization of cultural intermediation is conceptualised as 'platformed tastemaking', made possible by social media's platforming of: 1) curation, 2)

guidance and 3) measurement. Consequently, drawing on this literature review, I devise a preliminary conceptual framework outlining the platformization of cultural intermediation in social media. While concluding this chapter, more specific research questions are derived from the conceptual framework.

3.2. Background: Platformization

Digital platforms are slowly becoming an important part of both research and everyday work. Today most valuable companies are built on platform-based business models, while traditional business and the industries within which they operate are going through so-called platformization (Constantinides, Henfridsson & Parker, 2018; Van Dijck et al., 2018), and cultural industries are not an exception to that (Nieborg and Poell, 2018). The concept of 'platformization' is used to describe the ongoing transformation that has been altering the traditional value chains and how the traditional companies and industries operate. It is defined as a strategy for operating multisided markets and connecting buyers and sellers without controlling or owning the products or services being sold (Constantinides, Henfridsson & Parker, 2018). In this vein, Nieborg and Poell (2018) conceptualise platformization as the penetration of economic and infrastructural extensions of online platforms into the web. This is considered to be affecting the production, distribution, and circulation of cultural content in ways that have far-reaching cultural implications.

Social media are highly cited examples of platform-based business models (de Reuver, Sørensen, and Basole, 2016; Evans and Schmalensee, 2016; Lu, Goldsmith, and Pagani, 2013; Parker et al., 2016). As epitomes of digital platforms, they are assumed to have significantly altered the interactions built around cultural content by plat-forming long-standing forms of social and cultural activity through a distinct logic underpinning their operations. (Alaimo and Kallinikos, 2017; Langlois and Elmer, 2013; van Dijck, 2013). Contemporary culture is considered to be predominantly created, curated, circulated, and consumed on these platforms (Beer, 2013) in ways that permit the reactions of consumers of cultural content to immediately feed back to its creators and curators in the form of 'datafied user feedback' (Nieborg and Poell (2018). This is argued to create a recursive feedback loop (Alaimo and Kallinikos, 2017; Parker et al., 2016), which is assumed to represent the distinct aspect of the contemporary 'intermediary logic' of digital platforms like social media (Langley and Leyshon, 2017). This indicates that social media are rewriting long-standing cultural intermediation processes. Understanding this transformation requires looking at the building blocks of platforms closely.

3.3. Conceptualizing Platforms

The concept of a platform has multidisciplinary roots and has been subjected to scholarly attention in different fields. This has resulted in variegated accounts on platforms without a unified understanding of the phenomenon. Different streams of research have separately informed the study of platforms and contributed to our understanding of the platformization. This section reviews key insights from the literature on platforms informed by different theoretical perspectives in order to inform the conceptual framework devised to guide this study.

3.3.1. Platforms: Multisided Markets Reliant on Network Effects

The conceptual framework that will guide this study first builds on the business and economics literature, which draws attention to the multisidedness of digital platforms as well as the increasingly dominant institutional configuration created by these novel market structures. This strand of research provides valuable insight on the economic mechanisms and the managerial strategies underlying platform markets. They help to explain the increasingly dominant position of platform companies, which are argued to exercise control over the relationships among the end-users as well as the users and the third parties that try to target them (Nieborg and Poell, 2018). From the perspective of economists, platforms are special type of markets. They are often referred to as two-sided or multi-sided markets that facilitate transactions between different types of customers (Rochet and Tirole, 2003). For this reason, they are conceptualised as "matchmakers" (Evans and Schmalensee, 2016), connecting buyers and sellers who would not be able to interact with each other otherwise.

Studies in this field tend to adopt a transactional perspective to analyse the relationships among platform holders and users. One key insight provided by this perspective is that, a digital platform is different than a product platform because it does not necessarily have physical assets in the form of infrastructure resources and/or it does not generate value through product sales (Constantinides et al., 2018). Rather, they facilitate the connection between supply and demand (Baldwin and Clark 2000; Gawer 2014) within a dynamic ecosystem and thus enable new forms of production and consumption of value (Parker et al., 2016).

This study draws on these insights, highlighting the ways in which platforms operate as matchmakers. It adopts the definition by Parker et al. (2016) in which a platform is conceptualised as: "a business based on enabling value creating interaction between external producers and consumers. The platform provides an open, participative infrastructure for these interactions and sets governance conditions for them. The platform's overarching purpose is to consummate matches among users and facilitate the exchange of goods, services, or social currency, thereby enabling value creation for all participants."

One of the most striking aspects of platform-based businesses or the so-called multisided markets is argued to be their reliance on network effects (Parker et al., 2016). The term 'network effects' (Shapiro & Varian, 1998; Evans, 2003; Rochet and Tirole, 2003; Parker and Van Alstyne; 2005) describes a distinct dynamic that creates value for those who participate in a network, as well as those who own or manage it. These effects are observed to be either direct or indirect and positive or negative (Eisenmann, Parker and Van Alstyne, 2006). Positive direct effects are described as the increased value a platform generates for all participants every time a new participant joins to the platform (Katz and Shapiro, 1986; Parker and Van Alstyne, 2005). This is known as 'positive direct network externality' (Shapiro and Varian, 1998) and indicates how each person has a positive impact on other people in a specific network. Platforms are also subject to indirect network effects due to their multisidedness (Rochet and Tirole, 2003): Actors joining one side of the platform indirectly affect the value perceived by the actors on the other side of the platform. For example, when a new user joins Facebook, the platform is assumed to become more valuable for other users. This, in turn, increases the value of the platform for a wide range of complementary third parties, such as advertisers, brands and/ or app developers.

Likewise, the more producers create new content to be consumed on platform, the more consumers are assumed to be attracted, and it ultimately bring in more advertisers. Nevertheless, the common view is that it is not sufficient to secure one side to attract the participants on the other side to sustain the running of a platform business. Both sides of the market need to be brought on board (Rochet and Tirole, 2003; Parker et al., 2016). One notable challenge in this process is the so-called 'chicken-and-egg problem' (Hagiu and Eisenmann 2007; Parker et al. 2016), which refers to the challenges related to which side of the market should be brought on board first to attract the other side.

Another distinct aspect of digital platforms is their supposed capacity to transform linear processes within the traditional value-chains (i.e. product or content development and distribution) into an ecosystem model (Parker et al., 2016) constituted of different platform participants: consumers, producers, and third- party actors (Jacobides et al., 2018). Parker et al. (2016) describes this as the transformation of pipe-line businesses into networked markets through the *internalisation* of the activities involved in linear value flows. In doing so, they set the governance principles that guide the relationships between different actors interacting within a platform ecosystem (Gawer, 2014).

Tiwana (2013) broadly defines governance with regard to formal and informal control mechanisms and incentive structures used in the platform to organize participation. Control mechanisms allow platforms impose rules that reward and punish behaviour and establish certain form of behaviours as norms on the platform (Evans and Schmalensee, 2007). These controls can be informal or formal and may be in the form of gatekeeping and/or metricising behaviour (Tiwana, 2013). A platform's governance also includes *incentive structures*. This stream of research has generated valuable insights on how platforms exhibit architectural and governance rules to balance platform control with the necessary incentives to create and facilitate value-creating interactions (de Reuver et al. 2017; Ghazawneh and Henfridsson 2013; Parker et al. 2016; Tiwana 2015; Parker et al., 2016). One noticeable gap in this literature, however, is the lack of guidance on how such incentives and control mechanisms are embedded in the platform architecture.

Tiwana et al. (2010, p. 677) define platform architecture as the "conceptual blueprint that describes how the ecosystem is partitioned into a relatively stable platform and a complementary set of modules that are encouraged to vary, and the design rules binding both." These design rules (Baldwin and Clark, 2000) then serve as the means through which the relations of participants within a platform ecosystem are governed. Incentive structures have been examined as part of the managerial strategies underlying multisided platform markets, such as how network effects allow platform holders to set pricing structures where one side of the market, the "money side," covers the costs of the other side, the "subsidy side" (Evans and Schmalensee, 2016). For example, end-user access to social media platforms such as Facebook is free and subsidized by the money generated by businesses that are charged for the targeted market opportunities provided by the platform. This is made possible through user data used to consummate matches (Parker et al., 2016) between the

money side and the subsidy side. Relatedly, a platform needs to keep its participants engaged by encouraging and motivating them for active participation in ways embedded in the organising logic of a platform, in ways materialised through its design (ibid.). Despite being an important aspect of platform governance in terms of how a platform controls behaviour and incentivises participation, the role of a platform' design is barely investigated in this literature in relation to this aspect.

Complementing insight from this body of knowledge, the conceptual framework devised for this study also draws on the literature that is focused on social media as prominent instances of digital platforms that distinctively rely on data. In the next section, I review the key insights from this stream of research, which critically engage with the implications of social media's reliance on data. Digital platforms are assumed to engage in strategic actions and exert control over platform participant. This is considered by Boudreau and Hagiu (2009: p. 170) as being "distorted away from pure value creation in the ecosystem, and toward actions that lead to higher platform profits". In this regard, critical insights help to illuminate sociocultural implications of platformization. These approaches enable discussion of the data reliance of social media platforms in relation to their business models and pave the way for opening up a critical discussion on the implications.

3.3.2. Platforms as Sociotechnical Artifacts

The previous section reviewed previous studies and highlighted the economic logic that defines platforms as institutional actors. There are, however, observed gaps in this literature, especially with respect to sociotechnical aspects of digital platforms. For instance, the dominant tendency in this literature is to treat platforms as relatively static objects (Gawer, 2014; Tiwana, 2014). Moreover, despite recognising platforms as matchmakers, this literature provides little insight on platform's matchmaking operations in terms of the role of data and the sociotechnical design of platforms in this process (Alaimo and Kallinikos, 2016; 2017). Relatedly, the role of platform design is barely discussed with respect to the ways in which it embeds control and incentive mechanisms used to organise user-platform participation. These mechanisms need to work in self-regulating ways due to the concerns related to scalability (Parker et al., 2016). However, these aspects are merely discussed by business scholars. These gaps are partially addressed by a complementary stream of research on platforms.

Some of these gaps addressed by Platform studies (Montfort and Bogost, 2007, 2009; Burgess et al., 2017; Apperley & Parikka, 2018; Plantin et al., 2018). This literature primarily focuses

on the material, computational, and infrastructural dimension of platforms. As a specific branch, platform studies are described as a family of approaches to digital media concerned the relationship between computing platforms and culture. According to Montfort and Bogost (2009), this approach "*looks at how social, economic, cultural, and other factors led platform designers to put together systems in particular ways. The approach recognizes that not only the user's experience, but also interface, form and function, code, and platform, are fully embedded in culture. When undertaking such a task, interrogations of technical detail become important indeed. Such questions are posed not for their geek value alone, but in order to shed light on the relationships between technology and culture*" (p.5).

Taking social media as the centre of their focus, studies contributing to this body of knowledge conceive of platforms as computational infrastructures underpinning contemporary socio-cultural interactions (Langlois and Elmer, 2013; Van Dijck, 2013; Gerlitz and Helmond, 2013; Helmond, 2015; Plantin et al., 2018; Alaimo and Kallinikos, 2016, 2017). To start with, scholars in this stream of research emphasise that social media are rapidly evolving digital artefacts enabling and constraining online forms of communication and expression through their business imperatives (Gillespie, 2010; Langlois and Elmer, 2013; van Dijck, 2013; Helmond, 2015). From this viewpoint, van Dijck (2013) conceptualizes social media platforms as "dynamic objects that are tweaked in response to their users' needs and their owners' objectives, but also in reaction to competing platforms and the larger technological and economic infrastructure through which they develop" (p. 7).

These studies focus attention on the implications of the complex relation between social media's platform business model and their sociotechnical configuration. The starting point for much of this research is the user-platform relationship. The primary concern is to understand how the sociotechnical features of social platforms prescribe certain actions in line with platform's business models. Located at the centre of these developments, social media are recognised and frequently cited as one of the most prominent instances of digital platforms.

3.4. Conceptualizing Social Media as Platforms

Social media are increasingly recognized and conceptualized as a special kind of digital platforms operating as data firms. It is argued that, they are harvesting and repurposing users' behavioural data in ways that afford the construction of targetable markets for their customers, such as advertisers (Fuchs, 2012; van Dijck & Poell, 2013; van Dijck, 2014).

Previous work on social media reveal how structured platform affordances preformat, process, and articulate end-user activities, in ways that allow social media to produce repurposable data about users that can be monetized (Van Dijck, 2013; Van Dijck and Poell, 2013; Kitchin, 2014a; Gerlitz and Helmond, 2013; Alaimo and Kallinikos, 2016, 2017). These developments, according to van Dijck (2014), render personal information and behavioural data the new currency to be paid in return for free services provided by online platforms. This has turned the behavioural data generated into a kind of invisible asset, processed mostly as separate from its original context and outside of people's awareness (ibid.).

Behavioural data generated on social media are seen as highly rich sources for targeted market opportunities for advertisers, marketing professionals, and other third parties due to the novel aspects this new form of social data is assumed to reveal about individuals *qua* consumers. Social media, in this regard, are conceptualized as 'data platforms' (Alaimo and Kallinikos, 2016; 2017) that extract value and make profit from the data generated on the platform. They are also conceived of as *matchmakers* because they facilitate participants' interactions by matching end-users with relevant content '*matching* their *tastes'*.

Previous studies indicate that social media's matchmaking capacity is made possible by a complex relationship between the frontend and backend of social media. Design choices structuring the user-interface shape the activities performed by the users in ways that enables the construction of their taste profiles (Alaimo and Kallinikos, 2016, 2017) to be used at the backend systems for matchmaking operations, such as constructing personalised recommendations (Alaimo, 2014; Bucher, 2012; Couldry and Kallinikos, 2017; van Dijck, 2013) in line with the economic rationalities of platform-based business models. Thus they fulfil their raison d'être, *matchmaking*, by rendering platform participants *relevant* to eachother (Gilliespie, 2014).

This matchmaking can be conceptualized as a novel form of the audiencemaking process through which a social media platform facilitates multisided interactions. It involves rendering platform users target audiences of personalized recommendations, which suggest content that is either produced by other users or advertisers. Alaimo and Kallinikos' (2016; 2017) work exposes the complexity of this process. They describe the construction of taste profiles as a process in which "opinions, trivial concerns, daily and ephemeral pursuits, dispositions and experiences of users are all recorded via the modalities and affordances of social media,

pooled together and computed on a continuous basis to construct constantly updatable and, often, marketable profiles of daily living" (Alaimo and Kallinikos, 2016, p.3). As a result of this process, individuals are rendered 'users' by recreating them in the database as new social objects. Through the construction of '*data representations* of their taste, individuals enter into knowledge domain as data objects that can be identified, known, and acted upon (Desrosieres 1998; Foucault 1970). Thus social media establish the quantitative context of reference against which the platform operates.

Social media are argued to owe much of their distinctiveness to this capacity to produce *commensurable* data in the course of facilitating interactions on the platform (ibid.). Commensuration is defined as the transformation of different qualities in a common metric to reduce and simplify disparate information into numbers that can easily be compared (Espeland and Stevens, 1998). It strips away the context and represents relationships in abstract terms using numbers. Thanks to this distinct capacity, social media are conceived of as infrastructuring the trading of 'user taste profiles', produced through artificially created forms of sociality. This, according to Alaimo and Kallinikos (2017), is what renders social media data platforms and marks its distinction: facilitating interactions around transactional activities, such as buying, selling, booking etc., found to be less important for social media than the engineering of user-platform participation in ways that help in the creation of their taste profiles.

This can be considered to be social media's unique way of 'audiencemaking' (Ettema and Whitney 1994), through which individuals using social media are manufactured as 'particular kinds of audiences' (Bermejo, 2009) and rendered 'institutionally effective' (Ettema and Whitney 1994) as 'commodities' to be sold (Smythe, 1981) to advertisers or other third parties. Making tastes knowable through these means significantly differs from the types of data produced on other digital platforms, which are built to facilitate monetary transactions between sellers and buyers (Alaimo and Kallinikos, 2017). It also differs from the previous ways of knowing consumers' or audiences' cultural tastes that relied on more traditional instruments to measure tastes (Ettema and Whitney 1994; Moor, 2012; Napoli 2011; Wright, 2015). such as surveys, sales records, rating systems etc., which rely on categories produced by experts to segment consumers and audiences, such as demographics or genre classifications (Aaltonen, 2011; Lena, 2012).

This new audiencemaking is enabled through the distinct relationship between the frontend

and backend of social media. Social media make audiences 'institutionally effective' by producing specific kinds of data about them through their prescriptive designs and codification of interactions (Alaimo and Kallinikos, 2017). Thus they are assumed to recreate the conditions of traditional forms of sociality. This is described as social media's capacity to 'plat-form sociality' (van Dijck, 2013; Langlois and Elmer, 2013) by the aforementioned relationship between the frontend and backend of social media platforms.

3.4.1. Plat-forming: The Link between the Frontend and Backend

Previous studies that illuminate the relation between the frontend features and the backend processes of social media platforms (Alaimo, 204; Bucher, 2012) highlighted that social media platforms shape user-platform participation in the frontend in ways that help producing data for backend operations (Alaimo and Kallinikos, 2016; Langlois and Elmer, 2013; van Dijck, 2013). These backend operations underpin matching users with relevant third parties, the so-called 'money-side' (Evans and Schmalensee, 2016) discussed previously. This technological prescription and codification of social interactions on social media is referred to as 'plat- formed sociality' (van Dijck, 2013), while the processes are described as 'platforming' (Langlois and Elmer, 2013).

These notions are proposed to advocate moving attention away from what is happening on the interface level (what users say and post as part of identity construction and presentation of self etc.) to how frontend user interactions are shaped and rendered computable through the production, storage, and processing of data (Alaimo and Kallinikos, 2016; 2017; Couldry and Kallinikos, 2017; Langlois and Elmer, 2013; van Dijck, 2013). In this regard, critical scholars invite attention to the complex relation between the design and data operations of social media platforms. One such relation concerns how measures are used by social media platforms as a strategy to control and shape norms of behaviour and incentivise users to participate in platform activity.

As highlighted by the previous work in the business and economics literature, these aspects are crucial parts of platform governance. However, how these are structured by the design and embodied by the platform are not studied much (Gerlitz and Lury, 2014; Lampe, 2014; Pellikka, 2014). For instance, social media metrics are argued to be distinctive design features that are employed in ways that would develop a state of reactive tendency. As distinct design choices, these metrics, in the form of structured platform actions, are argued to be

strategically used by social media platforms based on the perceived ability of a particular metric to increase or decrease user participation (Grosser, 2014). In this regard, they are conceived of as intentionally used to invite and exploit 'reactivity' from users (Gerlitz and Lury, 2014).

The theory of reactivity suggests that measurement intervenes in the social worlds it depicts, causing people to think and act differently (Espeland and Stevens, 2008; p. 412). This is because people are reactive; that is, they change their behaviour in reaction to being evaluated, observed, or measured. From this perspective, Gerlitz and Lury (2014) argue that a distinct form of reactivity is purposefully created in social media to incentivise users to continuously participate in content production and platform activity. The main reason behind this is that social media platforms are essentially empty lots until users produce content and interact with each other through user-generated content. They are regarded as 'incomplete products' or 'empty frames', purposefully designed to be filled in in particular ways (van Dijck, 2013).

It is argued that social media users are directed to continuously produce new content and engage more with the platform due to the inequality that exists in most social media sites, whereby a small percentage of users produce or provide the majority of the content for a larger group of consumers (Pellikka, 2014). In this regard, motivating users to produce content is observed to be a common concern for most platforms. Accordingly, gamification, defined as the use of game elements in non-game contexts (Deterding et al., 2011), is found to be a commonly used strategy to create motivation (Lampe, 2014; Pellikka, 2014). It is described as the creation of self-regulating and self-reinforcing rules-based systems that incorporate 'feedback' and 'interaction mechanisms' with the goal of 'supporting the user's overall value creation by providing gameful experiences' (Hamari and Koivisto, 2013 p.3). It is observed that this is widely used to in social media to provide feedback in ways that trigger intrinsic motivation for users to engage in particular types of behaviours *desired* and *expected* by the designers (Lampe, 2014).

One of the primary reasons behind these efforts is to create, sustain, and expand upon network effects. Supposedly, a platform needs to find ways to continuously encourage users to produce content because how a platform scales and creates networks effects significantly depends on user-generated content (Parker et al., 2016). Constant user activity, on the other hand, is the primary data-producing force (Couldary and Kallinikos, 2017) for matching

users' with relevant content. At the centre of these operations lie the structured design features of social media used to organise user-platform participation. Preformatted actions metricise user reactions to create constant activity (Gerlitz and Lury; 2014), while codifying and encoding user activity to generate data influx to feed backend operations (Alaimo and Kallinikos, 2016). This indicates that social media are non-neutral mediators prescribing certain kinds of actions through user interface to increase platform activity at the frontend while generating data for the backend matchmaking process.

This relationship between the frontend and backend of social media is implicated in the ways in which cultural content is produced, curated, framed, circulated, and consumed. Similar to the ways in which they platform sociality, social media are thus argued to platform cultural production and consumption (Nieborg and Poell, 2018). Matchmaking operations sit at the centre of this platforming, yet there are very few studies that closely look at the aforementioned relationship in relation to the platformization of the processes mediating the space between the cultural production and consumption.

This relation underlies the ways in which social media create knowledge about their users taste to produce personalised recommendation at the backend, which in turn shape what users see on the frontend. This indicates that, in the course of plat-forming cultural production and consumption, social media also plat-form cultural intermediation. They are involved in the framing and presentation of cultural content and hence participate in tastemaking (Smith and Maguire 2014). Nevertheless, how exactly a social media platform performs matchmaking is either taken for granted or left as a black-boxed procedure in the literature.

This leaves a theoretical gap in our understanding of social media platforms and how they are involved in transforming traditional forms of cultural intermediation and, by extension, the tastemaking function of cultural intermediaries. To fill these gaps, Couldry and Kallinikos (2017) call for empirical research that deconstructs the logic underpinning the operation of social media platforms and the sociotechnical features shaping the terms of user platform participation. They argue that this necessitates "*prolonged immersion into particular research contexts (i.e., intensive case studies) that may provide the opportunity to unravel the design choices, technological systems and economic rationalities of social media platforms*" (ibid., p.19). In this light, the next section is dedicated to developing a conceptual framework to outline the platformization of cultural intermediation. The purpose is to unpack the

complex relationship between the frontend and backend of a social media platform in relation to their matchmaking operations.

3.4.2. Plat-formed Intermediation

Previously reviewed economic and management studies also recognise digital platforms' intermediary role as an emergence of a new form of intermediation (Gawer & Cusumano, 2002; 2008). As an extension of their matchmaking capacity, platforms are also conceived of as intermediaries enabling 'value-creating interactions' between external producers and consumers (Parker et al. 2016) in ways that structure 'market encounters' in digital space (Çalişkan and Callon, 2010: 14-16). Platform's intermediation is conceived to be distinctive because, integral to their capacity to structure encounters in the digital space, platforms enable connectivity within the multisided markets while coordinating the network effects of this connectivity (van Dijck, 2013). The distinction of this new form of intermediation is also attributed to the ways in which platforms disrupt the traditional value chains and the long-standing separation between creators, content producers distributors, and consumers (Benghozi, 2016). Platformization of intermediation is considered to pave the way for a profound redefinition of the notion of intermediation, which, according to Benghozi (2016), requires considering platforms as the "active third parties operating in parallel with producers and consumers, either to structure the supply of products or services, or to take over part of the consumer's decision making responsibility" (p.5).

Langley and Leyshon (2017) describe this as a 'distinctive intermediary logic'. They argue that this novel intermediary logic materialises through: 1) the contingent configuration of platforms as discrete modes of socio-technical intermediaries and business arrangements, and 2) using data analytics to prescribe, structure and regulate users interactions brought on to the platform through the rhetoric of the so-called 'participatory culture' or 'value co-creation'. Consequently, the processes in the traditional value flows in cultural industries that are involved in the production, curation, circulation, and consumption of cultural content get plat-formed (Nieborg and Poell, 2018).

A platform's intermediary logic is argued to be in sharp contrast with the traditional 'editorial logic', which more explicitly relies on the 'choices of experts' (Gillespie, 2014:192), such as the curatorial practices of gatekeepers or the recommendations and guidance provided by critics and reviewers (Wright, 2015). These cultural intermediaries used to acquire their legitimacy from their knowledge and expertise in a particular field (Smith-Maguire and

Matthews, 2014; Wright, 2015). On the contrary, this new intermediary logic result in new forms of curation and guidance that are built on novel forms of knowledge created about users' tastes. This new logic also owes its distinctiveness to the ways in which platforms acquire *legitimacy* as trusted intermediaries based on their ability to create relevance among platform participants through the knowledge they create about users' taste.

Categorising, Measuring and Making Up People

This is closely linked with long-standing trust in the objectivity of numbers (Porter, 1995) as well as assumptions regarding the nature of data. Previous work in the literature concerned with the implications of classification, quantification, and measurement showed that data are usually seen as *pre-analytical* or *pre-factual* (Kitchin, 2014b). Henceforth they are assumed to represent what exists prior to interpretation. In line with these assumptions, data are considered to be the 'raw material' from which information and knowledge are built to 'manufacture audiences' (Aaltonen, 2011). They are viewed as *representative*; that is, they afford to capture the world as it is in the form of numbers, characters, symbols, images, sounds, electromagnetic waves, bits, etc. (Floridi 2008, 2010; Rosenberg, 2013). Thus they are assumed to take the so-called 'facts' out there and render them abstract, discrete, aggregative, and meaningful.

These assumptions are challenged by some scholars who emphasize that data are never simply neutral, objective, independent, raw representations of the world, but rather do active work in the world, as they are framed in particular ways to achieve certain aims and goals (Dalton and Thatcher, 2014; Kitchin, 2014b). They are considered to be the products of particular assumptions, techniques, and technologies involved in the production and processing of data (Bowker and Star 1999; Lauriault, 2012; Kitchin 2014a). In this regard, as Gitelman and Jackson (2013: p.2) succinctly put, "raw data is an oxymoron"; "data are always already 'cooked'." This cooking is not free of consequences. Hacking's works variously showed the implications of the process of producing and legitimating knowledge on human behaviour by revealing the generative aspects of these processes.

The notion of a 'looping effect' (Hacking 1999, 2007) concerns how data are classified and organised; how a distinct form of data ontology comes into existence; and how it reshapes that which has been classified. Hacking explains this mechanism as a form of 'dynamic nominalism', which suggests that there is an interaction between data and what they represent. This, according to Hacking (1986, 2007), leads to mutual changes and results in a

process of ‘kind-making’, which is he refers to as ‘making people up’. The ‘kind’ in this process may be kinds of people, kinds of behaviour, kinds of action, and/or kinds of experience (Hacking 1986), which at the end of this process become ‘moving targets’; classifications affect the classified subjects and, in turn, the reactions of the classified affect the classifications themselves. Consequently, this loop-like process results in rendering people ‘moving targets’.

In Hacking’s words, “What was known about people of a kind may become false because people of that kind have *changed* in virtue of how they have been classified, what they believe about themselves, or because of how they have been treated as so classified. There is a ‘looping effect’” (Hacking, 1999: p. 104, emphasis added). ‘Making up people’ is to a large extend made possible by what Hacking (2007) calls the ‘engines of discovery.’ These engines usually take the form of the statistical analysis and the underlying techniques and technologies involved in investigating human behaviour. These tools are argued to reshape society and individuals in the form of data ontology, therefore considered as ‘engines for making up people’ (ibid.).

This process becomes more complex when categorization becomes computerized in social media platforms. It is argued that, categorization work naturalized in social media through the solidification of the categorised that are used to represent individuals in the form of data (Couldry and Kallinikos, 2017). As users take up the personalised recommendations that are produced by these data, the relation between categories and the objects/people they categorize are assumed to be further reinforced and become naturalised over time. This is considered to result in category-naturalization as the categories and the categorised objects ‘converge’ and become ‘mutually constituted’ (Bowker and Star, 1999). These aspects of social media’s data processing are also closely related with the productivity (Beer, 2015) and performativity of measures (Espeland and Stevens, 2008), ‘co-constituting’ (Gerlitz and Lury, 2014) the phenomenon they purport to describe.

3.5. Conceptual Framework

In light of the key insights provided by the reviewed literature, this study conceives of social media as data ‘platforms’ operating as multisided markets and reliant on the network effect: increased participation in one side of the platform directly or indirectly affects the other sides of the market. Due to this dynamic, social media platforms employ various strategies to control and incentivise user-platform participation with the aim of keeping users active

through their prescriptive design and backend operations. Closely related to this, social media are also conceptualised as 'matchmakers': they bring together and consummate matches among users as well as between users and third parties, which are interested in reaching out to consumers for targeted market opportunities. This is assumed to be a novel form of audiencemaking. As highlighted by previous works, social media perform matchmaking by crafting taste profiles for their users by structuring and encoding user-platform at the *frontend*. They do so in line with a distinct logic underpinning the platform design and *backend* data operations. In this regard, social media's matchmaking is accepted here to be the result of the relationship between the frontend and backend of the platform. This relationship shapes both the user-platform participation and the ways in which users are matched with the relevant content. Thus social media is involved in the determination of what content become available for consumption and how it is framed and presented to its audience.

This indicates that social media are transforming the cultural intermediation process. They involve in framing and presentation of cultural content and guide audiences by measuring their taste and curating content on the basis of the computed relevance between the two. Explaining the contours of this transformation requires deconstructing the relationship between the frontend and backend of social media in ways that illuminate how a social media platform creates data representations of tastes and use them during matchmaking operations. To this end, a preliminary conceptual framework is devised to outline how social media operate as matchmakers and how this is implicated in the platformization of cultural intermediation. This section outlines the main constructs of the conceptual framework and describes their relationship.

3.5.1. Constructs

To enable this research, the conceptual framework is informed by two bodies of literature aiming to systematize the exploration of social media platforms as matchmakers. First, it is informed by the previous works on platforms that shed light on the ways in which social media afford the *Platformization* of traditional value chains, which entails the internalization of the entire end-to-end flow we find in traditional value chains. To unpack this process, this study deconstructs social media platforms' design logic. This is to explain how social media's design enables disassembling users' dispositions and turns them into data, which are assumed to represent their tastes. Complementary to this, the framework is informed by the literature

on *Cultural Intermediation* and conceptualizes the curation of content, guidance of audience and measurement of taste as its prominent instances.

3.5.2. Relationship between Constructs

In order to explain the platformization of Cultural Intermediation, main constructs are operationalized using their defining attributes. *Platformization* is operationalized as the ways in which a social media platform *internalises* a value flow. The concept of cultural intermediation is operationalized as constituted of: *curation* (what gets available), *guidance* (quality assessment and recommendation), and *measurement* (how taste made known), which were previously identified as the key instances of the cultural intermediation processes accommodated in more traditional linear value flows. Consequently, the 'platformization of cultural intermediation' is outlined as a social media platform's internalisation of these processes through its distinct sociotechnical configuration. Consequently, the platformization of cultural intermediation can be outlined as follows:

A social media platform *internalizes* curation, guidance and measurement activities through its distinct design logic. This logic entails bringing together content producers and consumers and facilitating their interactions through a core interaction (Parker et al., 2016), and organizing user-platform participation at the frontend in ways underpinning by backend data operations (Alaimo and Kallinikos, 2017). In the course of this internalization, the processes mediating between the production and consumption of cultural content are replaced by the core interaction of a social media platform. The platform facilitates user interactions by performing multisided matchmaking, which involves a novel form of the audiencemaking process constituted of encoding, aggregation and computing. These matchmaking operations are underpinned by data influx, which necessitates continuous user activity on the platform. This is accomplished by embedding governance strategies within the platform design in ways that control and incentivize activity in self-regulating manners. Thus a platform architects governance strategies to keep users engaged and facilitates their interactions by consummating matches through backend data operations.

In the following section, I describe in detail how the main constructs are conceptualized in this study.

i. *Platform Design*

Platformization relies on a distinct design logic that entails internalising the traditional value flows on the platform in the course of the development of a digital platform (Parker et al., 2016). This is accomplished by digitising the two ends and the processes mediating between these two in ways that produce data about the participants based on the interactions that take place on the platform (ibid.). In this vein, the term 'platform design' is defined here as the 'overall logic' underpinning the *development* of a digital platform. This logic is conceived of as building on particular assumptions and decisions regarding the roles played by the participants and the rules prescribed by the platform to organize user-platform participation. There are two main roles for participants involved in building a platform and the supporting interactive ecosystem around that platform: consumers and producers. In this basic setup, the platform connects the demand and the supply sides by facilitating participants' interactions through consummating the match between consumers and producers. In the following, I outline the core design principles and data operations underpinning the development of a social media platform.

ii. *Core Design Principles*

Platforms internalise the entire end-to-end flow of linear value flows in ways that generate data about different aspects of this very flow itself. This requires identifying, digitizing, and internalising the 'source' and 'destination' of the flow. Thus the two-ends of a value flow, the points of production and consumption, get digitized in ways that also generate data about the processes mediating between these two ends. Parker et al.'s (2016: p.5) definition of a 'platform' encompasses three aspects of the overall design logic underpinning the development of a platform: Platforms: 1) enable interactions between users constituted of external producers and consumers, 2) provide an infrastructure that embodies the rules of participation, and 3) perform matchmaking between users to facilitate their interactions. In addition, the development of a platform requires the consideration of factors such as: how to invite and sustain participation and engagement, and how to provide tools and services that make it easy for producers and consumers to interact in ways that scale the platform rapidly through positive network effects. In order to deconstruct this basic design logic, Parker at al. (2016) focus on the fundamentals of platform design, scrutinizing 'what exactly a platform does', and 'how exactly it works' and proposing a framework delineating the main design principles and key platform functions underpinning the development of a platform.

- **The WHY of the Platform**

According to this framework, the design of a platform significantly relies on 'the core interaction'. It is defined as the single most important form of activity that takes place on a platform. It is argued that, irrespective of the type of platform, the design of the core interaction follows a common template constituted of three main activities performed by the 'producers' and 'consumers': 1) creation, 2) curation, and 3) consumption of value units. All the actions permitted on a digital platform are assumed to boil down to one of these three activity types.

- **The HOW of Platform Design**

This framework also outlines three key platform functions underpinning the ways in which platforms scale through network effects: 1) pull, 2) facilitate, and 3) match. The pull function involves the ways in which consumers and producers are incentivized and motivated for continuous participation. A second important function is facilitating interactions between the participants by providing them with an infrastructure that lays out the principles governing the interactions among the users. Lastly, platforms consummate matches between users using the data generated about the users during their interactions.

iii. *The role of 'Data' in Platform Design and Functioning*

I have discussed how Parker et al.'s (2016) framework helps to understand the basic design logic underpinning the development of an online platform. Their framework suggests that platforms facilitate interactions between producers and consumers around a core interaction by matching users. As pointed out previously, data are a significant element of a platform's matchmaking function. However, despite outlining matchmaking as a crucial platform function, Parker et al. (2016) do not explain how exactly data is produced and used in these operations. This limitation is overcome in this study by complementing their insights with the framework proposed by Alaimo and Kallinikos (2017).

Alaimo and Kallinikos' (2017) work illuminates the contours of backend data operations underpinning the production of personalized recommendations, which is the key aspect of the matchmaking function. Their framework outlines the logic through which social media platforms organize, record, and measure user platform participation in the frontend in ways that generate data for the backend. According to this model, social media platforms shape user-platform participation through three fundamental data operations: 1) encoding, 2)

aggregation, 3) computation. Social media's unique interface design enables the crafting of user taste profiles out of structured and stylized actions, such as: tagging, 'liking', commenting, reviewing, and so on. These actions enable the facilitation of user interactions in ways that allow the encoding of user-platform participation. The platform encodes user-platform participation in ways that translate it into discrete data tokens.

Encoding renders user behaviors into data representations, which are taken to be indicators of their dispositions. Resultant encoded data are aggregated to form bigger entities. This is accomplished by piling-up the singular platform activities that were encoded into data. Aggregation is a distinct instance of abstraction. It helps to establish, out of the singular actions of individual persons, new social objects. Thus real persons are rendered data and become platform users, defined in the database through their actions. Encoded and aggregated user behaviours are then used to compute users taste to match them with relevant items through personalized recommendations. Thus computed, data representations of users tastes allow the platform to make connections among the users and content in the so-called personalised ways.

3.6. Research Questions

In light of the conceptual perspectives introduced in this chapter, the main argument is structured around a refined research question and a set of sub-questions to better operationalize it for a focused analysis of the empirical case study.

RQ: How does a social media platform internalise cultural intermediation?

1) How is content organized (curation)?

2) How are users assisted for discovery (guidance)?

3) How is knowledge on taste is produced (measurement)?

3.7. Conclusion

In this chapter, I reviewed the literature, staged a dialogue between different bodies of scholarly work, and devised a preliminary framework to guide the analysis of the case study. Previous work on platforms in the management and economics literature informed the delineation of a distinct design logic underpinning the development of a digital platform. This is combined with a complementary theoretical lens that perceives digital platforms as sociotechnical entities. Taking social media as distinct instances of digital platforms, this

strand of research highlights the ways in which platforms organise user-platform participation in prescriptive ways in line with the economic rationale underpinning their operations. Finally, these developments are discussed in light of the literature concerned with the implications of data processing operations that rely on measurement, categorization and abstraction.

Each of these research traditions offers a valuable avenue to analyse different aspects of digital platforms, with social media as highly cited examples of these complex organizations. Drawing on the key insights from these scholarly accounts, I recognize how digital platforms transform linear value flows into complex market structures. This transformation is referred to as platformization and conceptualized as the reorganization of the processes involved in the traditional value chains into multisided markets through their internalisation by a digital platform. Complementary to this**,** I focus attention on how platforms shape interactions at the frontend to generate data for the backend operations. Thus I have highlighted how platforms operate as matchmakers. These perspectives are thus combined in a framework to outline the platformization of cultural intermediation. Consequently a conceptual framework is devised to guide the analysis of the empirical findings and a more refined set of research questions are presented.

4. Research Design

4.1. Introduction

In the previous chapter, the devised conceptual framework guided the refinement of the broad research question and the formulation of specific sub-questions. This chapter discusses the research design used to address these research questions. It is worthwhile to note here that this study adopts a theory-driven research design rather than theory testing. It aims to develop a theory to explain platformization of cultural intermediation. This is accomplished through a qualitative methodology based on a single explanatory case study. It includes an initial exploratory phase, which helped to develop an understanding of the history and context of the investigated company and its platform-product. This phase helped to delineate the contours of the empirical investigation, thereby shaping and guiding the rest of the case study. In the explanatory phase, analysis evolved hand in hand with the data collection through the ongoing interpretation of research findings. This helped in refining the analysis, and it resulted in the final explanatory framework presented in the analysis chapter.

This chapter is structured as follows: First, I discuss the ontological and epistemological stance of the research project. Following this, I discuss case study as the chosen methodology and justify the adoption of a qualitative approach. Then I explain the rationale behind the selection of music as the taste case and the reasons behind the selection of the particular case chosen as the object of this research project. I conclude this section by delineating the case study design, which has gradually taken a form that is constituted of an initial exploratory phase, followed by a more in-depth explanatory phase. In the next section, I explain corpus construction (Bauer and Aarts, 2000) as the chosen technique for data collection and selection. The subsequent section explains the data analysis phase. Finally, I conclude the chapter with a brief summary.

4.2. Epistemology and Ontology

This study adopts a constructivist approach in combination with a critical realist ontology. In particular, Ian Hacking's (1999) approach, referred to as the 'construction of kinds', is adopted. Following Hacking, the objective of this research project is to explain the particularities of the *kind* of data representations social media platforms produce about individuals in the process of platforming cultural intermediation and, more notably, in the

course of producing knowledge about cultural taste.

I acknowledge that the phenomenon under investigation entails some aspects that are not easily *observable* (i.e. the rationale behind the design choices and data operations as well as the assumptions underpinning them). This affects how we understand the ways in which social media platforms mediate contemporary living. However, how exactly they take over the traditional cultural intermediaries' roles (i.e., gatekeepers, reviewers and critics) cannot be understood by *merely* studying the human-technology interface of social media platforms because frontend user practices are heavily prescribed and scripted by the platform design and recursively shaped by backend data operations. In this regard, understanding social media's mediation requires looking beyond the human-technology interface (Kallinikos, 2004), and examining how their sociotechnical configuration and underpinning data operations shape what users see on the interface. Moreover, digital artifacts are intimately connected with the wider digital ecosystem in which they realize their utility (Kallinikos 2006; Kallinikos et al., 2013). The operations of digital technology extend far deeper than the human technology interface and cannot be sufficiently explained by interpretation of local agents (Kallinikos, 2004). This study therefore adopts a constructivist epistemology in combination with a critical realist ontological stance that engages in explaining the unobservable mechanisms that produce the observed events. I will now elaborate on these points in detail.

4.2.1. Epistemology

Epistemology is concerned with how researchers know what they know. It shapes the assumptions underlying the scientific approach adopted for research. It also determines to what extent the phenomenon can be known and how it can be researched (Crotty, 1998). This study adopts a constructivist approach, assuming that knowledge and representations of reality are actively constructed in social processes. Social constructivism is a very broad perspective that encompasses many variants of how the social construction of our realities happen, ranging from the inevitability assumptions of Hacking, to a more classical view of objective and subjective realities by Berger and Luckmann, to the extreme constructivism found in early Latour (Alvesson and Sköldberg, 2009). In this fragmented picture, I adopt the view that suggests that there is a reality 'out there' (Sismondo, 1993). That is, I recognize the distinction between a subjective and an objective reality – a 'reality' (in scare quotes) that is subjective and a reality (without quotes) that is objective and "cannot be wished away"

(Berger and Luckmann, 1991). But beyond this distinction, I adopt a particular approach to constructivism, which Ian Hacking (1999) refers to as a 'construction of kinds'.

There are implicit assumptions behind the main research question of this research. These concern how social media platforms transform cultural intermediation in the course of internalising the traditional related processes through the relationship between the frontend design and backend data operations. The core assumption is this: if social media platforms are designed differently and supported by different operations than the current ones, then we could have had a different configuration for contemporary forms of cultural intermediation than what we have today. This suggests that social media platforms transform this process in *particular* ways. But there is awareness that things could have been different. Therefore, what we have today is *not* an *inevitable* result of technological developments or social processes. Rather, it is a consequence of particular choices, conventions, assumptions, arrangements, practices, standards and other historical and contextual contingencies.

Hacking (1999) argues that we often take for granted many of the assumptions that come to shape social reality, although most of the social phenomena exist in the way they do now (i.e. economy) as a result of particular, albeit contingent, definitions and choices. In this regard, Hacking is against *inevitability*, and, hence, against *objectivism*. He instead emphasizes that things could have been different if they were framed, discussed and acted upon differently. Therefore, his constructivist approach starts with the idea that there is a social world and it is taken for granted. What is under construction in this picture according to Hacking is not necessarily *the reality,* but its *kinds*. He consequently advocates focus on how is it that a particular social reality is constructed, made neutral and natural based on specific assumptions about the kind of reality it refers to. Many things in social life are indeed constructed, but it is crucial to bear in mind that their current form is not inevitable. Hence his constructivist approach necessitates looking beyond the constructed aspects of things and explaining the ways in which they become 'taken for granted.' Adopting this approach, this study aims to explain the 'data-based' forms that prevailing cultural processes take within the context of social media and how these are taken for granted.

4.2.2. Ontology

Ontologically, this study adopts a critical realist stance to understand and explain unobservable mechanisms that produce the observable aspects of the phenomenon under

study (Alvesson and Sköldberg, 2009). It makes a distinction between the real, the actual, and the empirical. The real is constituted of the enduring structures and mechanisms generating causal effects (Sayer, 1992). Social structures influential on the consumption side of tastemaking (i.e the socio-economic conditions and class-specific habitus of individuals) have long been recognized as shaping how taste is formed (Bourdieu, 1984). In the same way, social structures shaping the production-side of the processes of tastemaking (i.e. creation, curation, guidance and measurement) significantly condition actions and interpretations, and they eventually become influential in the making of taste. Thus, even if not immediately observable, or even if not observable at all, these structures generate actual events by interacting with each other. This may result in the production of empirical events and become observable. Departing from these basic assumptions, this study investigates observable events in the chosen case study in order to identify the mechanisms producing them. Thus its purpose is to explain the kind of conditioning that social media platforms generate.

4.2.3. The Compatibility of Epistemology and Ontology

An ontological stance of realism and a constructivist epistemology are known to be quite compatible (Crotty, 1998). This is because, as Sismondo (1993) suggests, recognition of the socially constructed aspects of reality does not necessitate a rejection of the idea that there is a reality 'out there.' This reality might be rendered meaningful or gain meaning only through social processes, but its realness is what limits and allows its social constructedness, as reality is layered: "To say that meaningful reality is socially constructed is not to say that it is not real (Crotty, 1998 p.63)." Thus constructivism and realism can be coherently combined. Their combination enables a researcher to move beyond what is observable at a superficial level and research the underlying unobservable aspects of a phenomenon. Given how this study investigates a complex socio-technical phenomenon – namely, how the configuration of social media platforms relates to a highly abstract topic like taste - this combination is considered to be suitable for this research project.

The ways in which social media platforms operate as cultural intermediaries cannot be empirically observable. For this reason, it should be explained through retroductive reasoning. The observation of events and analytical reconstruction of the structural relation between them paves the way for theorizing on hypothetical mechanisms (Sayer, 1992). In this vein, this research project explores how cultural intermediation is transformed through platformization by investigating how this process is internalised by a social media platform.

Here, cultural intermediation processes are studied as *real* entities, while their relations are assumed to create the structural conditions influential in shaping cultural tastes. Consequently, the following hypothetical mechanisms are proposed:

- There are certain processes of mediation between the production and consumption of cultural content, the prominent instances of which are: i) curation, ii) guidance, and iii) measurement.
- Social media 'plat-form' these processes by *internalising* them through their platform systems, which operate as a matchmaker.
- The ways in which users' *taste profiles* are constructed to be used in social media's matchmaking operations are the observable 'events'. Hypothetically, they are created through the distinct relationship between social media's frontend design and backend data operations underpinning their matchmaking operations. This constitutes the basic object of this research.

The analysis and discussion are performed based on these assumptions. In this case, the observable events are the user taste profiles constructed by the platform. That is because the researcher can collect data from the informants about the processes in which these profiles are produced. These profiles are also produced based on presupposed conventions regarding what might be indicators of one's taste. It is the role of the researcher to identify such underlying assumptions, to synthesise them with the actual ways in which taste profiles are constructed and, in doing so, to deconstruct the relationship between the frontend and backend of social media in ways that can explain the platformization of cultural intermediation.

4.3. Methodology

4.3.1. Case Study Research Strategy

Case study research is chosen and employed as the research strategy in line with the objective of this research project. According to Yin (2003), case study is the suitable methodology to research "a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin, 2003, p. 13). Given how this thesis is interested in exploring, understanding and explaining developments in a research domain that is not well charted through an established theory, it necessitates in-

depth scrutiny of the phenomenon in its own context. Thus a single explanatory case study has been conducted.

The need to account for and understand issues and concerns involved in the development of a social media platform that are not easily observable arises from the 'how' questions (Yin, 2003) posed by this study. This study departs from two basic questions: "how do users see what they see on the interface level of social media platforms?" and "how is this reconditioning the traditional ways in which individuals encounter cultural?" and so on. To answer such 'how' questions, one must scrutinise various issues in-depth and answer further questions, such as: "How does a social media platform construct taste profiles of users?"; "How does a social media platform identify the type of actions that users take on the platform, which can then be taken as the basis behaviour signifying their taste?"; "How does a social media platform represent users tastes?"; "How are users supposed to use the system, or to what extent does the system allow flexibility in the ways users provide input data?"; "How is content curated?"; "How are personalized recommendation lists composed and how they are presented to users?".

Yin (2003) argues that having a conceptual framework is one important difference between case study research and related methods such as ethnography or grounded theory. The latter approaches deliberately avoid specifying conceptual frameworks prior to data collection and encourage researchers to proceed into data collection as soon as possible. Case study, in contrast, starts with a guiding conceptual framework, which helps the researcher to avoid difficulty and confusion in the course of data selection and hence eases data collection. Moreover, case study as a design choice allows one to investigate a research domain that is not well understood by using theoretical propositions even in the lack of an established theory charting the research territory (Yin, 2003).

This helps the researcher to point attention, to limit scope, and to suggest possible links between the observed phenomena. It provides a logic that helps to link propositions to the data to be collected and then to its analysis. It also allows researchers to change and/or refine their initial theoretical propositions and search for new ones. Thus they can continuously elaborate on their research questions, theoretical propositions, units of analysis, and criteria used for interpretation, as well as the links between data and theory.

In light of these points, this research is conducted based on a single case study with a single

unit of analysis: that is, a social media platform on music discovery. More specifically, the object of study is a platform as a digital artifact operating on the web. In this regard, the single unit of analysis is comprised of: the platform design, data-based operations, and the relation between the two; the processes in which these are designed, developed and linked to each other; and the assumptions underpinning these. The case study is justified as a research design choice on the basis that it allows a researcher to use multiple sources of evidence, including documents, interviews, and observations, as well as artifacts - an example of which, in this case, is the platform itself.

4.3.2. Qualitative Research

This study adopts a qualitative approach because the research domain is a cross-disciplinary area that is based on different strands of literature. It is not yet sufficiently explored and explained with established theories. Studies focusing on the intersection of culture and technology are usually limited to investigations concerning the consumption side (i.e., user practices taking place at the level of user interface), and are mostly concerned with how individuals make sense of and enact technology. Issues concerning the production side are studied less, and if/when recognized, technological reconditioning is often attributed to the role of proprietary algorithms. This is an approach problematized by this study, as such studies fail to provide an in-depth explanation for the inner workings of sociotechnical arrangements like social media platforms and tend to black-box the role of the platform by mystifying the role of algorithms.

Social media's sociotechnical configuration is substantially involved in the remaking of everyday cultural experiences. For this reason, it is a complex research domain that requires in-depth investigation (Couldry and Kallinikos, 2017). As the body of knowledge in this domain is still in the making, a research project on this topic requires theory development rather than theory testing. Thus a qualitative inquiry is more suitable for this research because it allows knowledge to be built gradually in a specific domain that is either unknown or too complicated to be tested with a theory based on a predetermined sampling technique.

Epistemological concerns, discussed in relation to the implications of techniques of making up people (i.e. classification, quantification, measurement), also inform the methodological considerations. Just like the ways in which classifications and measures play crucial roles in making up (or constructing the reality about) what is classified and/or measured (Espeland

and Stevens, 2008; Hacking, 1986), empirical investigation can also be considered as a process of constructing data for analysis. That is because what is observed is already shaped by the selectivity of the observer. In a way, the researcher co-constructs the phenomenon under investigation, initially through the questions she poses, as well as the assumptions underpinning those questions. Likewise, the research strategy used to answer those questions significantly shapes the kind of answers that can be given to the posed questions. In the same vein, criteria used in the selection of the units to be analysed, as well as the analytical and theoretical lenses used to guide the collection and analysis of data, all play role in the resultant research output.

Methodology works self-referentially, as it aims to reduce the number and refine the form of potentially infinite ways of seeing and observing the research domain (Müller, 2008). In this way, it is noteworthy to bear in mind that the procedures and the selection criteria employed in qualitative research often have cyclical aspects. Employing qualitative methods in theoretically informed ways with the aim of developing a theory can be considered as an exercise in which an unknown territory is investigated with the help of known theoretical constructs (Bauer and Aarts, 2000). In this regard, qualitative research is considered to be a process in which an unknown or a not-well-charted phenomenon is gradually described or explained through purposive sampling. This process resembles 'looping effects' (Hacking 1999, 2007), wherein the output of a process also becomes the subsequent input feeding back to that same process. It is in this regard that this research gives particular importance to the issue of self-reflexivity, and it acknowledges the process of *data collection* also as a form of *data construction*.

This research project emphasises that methods, as techniques or tools of measurement used in investigating a phenomenon, *do participate* in its definition, and thereby they attend to the making of the very phenomenon under scrutiny. Hence, according to Law (2004, p.143), method is *not* "a more or less successful set of procedures for reporting on a given reality." It is rather *performative*: "It helps to produce realities." For this reason Müller (2008) asserts that, what is accounted for and regarded and respected as 'scientific truth' can also be seen as an achievement of the very science investigating that 'truth.' A self-reflexive stance necessitates an acknowledgement of this. In this regard, this study does not claim to account for an objective reality. Instead, the aim is to construct informed ways of seeing the research domain through structurally determined ways of seeing.

4.3.3. Theory Building and Generalisability

This study's aim is to build a theory that is able to explain the platformization of cultural intermediation by social media. However, due to the lack of an established theory in this domain, it also entails an exploratory phase, after which more new themes are expected to emerge based on empirical evidence. These will be complementary to the initial concepts identified to guide this study. For this reason, the corpus construction technique (Bauer and Aarts, 2000) is preferred as the data collection method. It enables the investigation of a not well-known domain with known categories through a cyclical process of selecting data, analysing it, and selecting again for further refinement of ideas for theory building.

This is connected to the issue of generalisability in research. Yin (2003) argues that case study researchers should aim toward analytic generalisation, rather than statistical generalisation, to explore the linkages of causes and effects in a particular case. It is a way of generalising toward a theory, not generalising toward some defined population that has been sampled, since findings may have much wider relevance than the particular case being studied. In this regard, the aim of this study is to arrive at analytical generalisations, rather than statistical ones, and to find the general from the specific. This type of generalisability involves generalising from constructs, concepts or variables to a theory. A preliminary version of theory or theoretical propositions can be formulated "based on the synthesis of ideas from a literature review (Lee and Baskerville, 2003)." These initial conceptual ideas can be developed and refined in light of data. This paves the way for theory development based on a hybrid approach informed by theory and empirical evidence.

4.3.4. Case Selection and Case Study Design

In this section, I outline the rationale behind selecting *music* as the cultural domain to explore cultural intermediation and, by extension, tastemaking. Then I proceed to explain why the studied case has been chosen.

i. Selecting Music Discovery as the Empirical Domain

Music is deeply interwoven into everyday living due to is ubiquitousness as an accompaniment to daily activities such as: running, studying, working, driving, dining, and so on. Other forms of media - such as books, film, or TV content - are not consumed as easily, as frequently, and as ubiquitously as music objects - such as individual songs, albums, or,

more recently, playlists. In this regard, digital music represents a mode of ubiquitous listening, which acknowledges that most listening happens "alongside or simultaneous with other activities" (Kassabian, 2001: p.15).

Music has been mobile ever since the introduction of the mobile technology called the 'walkman', and, after the widespread adoption of smart devices, listening became more ubiquitous than ever before. Smart devices have provided listeners with the opportunity to have constant access to online sources that host vast libraries for music. In such a setting, music is streamed via a mobile device and its Internet connection; listeners do not need to own music. Thus music is now offered as a *service* –a notion referred to as Music-as-a-Service (MaaS)[2]. As users turn on their music players and go about other activities, the sound started to function as a soundtrack to other activities, rather than being a specific activity on its own. Consequently, the omnipresence of music is assumed to create a fabric that accompanies the patterns of listeners' everyday actions. If music is the soundtrack of daily lives, then the notion of music as a service, making it available anytime and anywhere, is considered to be a new mediator in listeners' relationship with music. That is because it renders access and listening even more ubiquitous (Kassabian, 2013). As a result, "the more ubiquitous music appears, the more difficult it is to conceive of it as a separate and distinct experience outside of everyday activities" (Morris, 2011). In light of these points, I picked music as a taste case to explore the transformation of cultural intermediation and, thus, the tastemaking process.

ii. Selecting the Case Study

Once the problem domain was identified and music was selected as the suitable taste case to investigate, I reached out to several industry professionals to get more information about the

[2] This notion is referred to as Music as a Service (MaaS). In line with the service dominant logic outlined by Vargo and Lusch (2004), MaaS refers to the shift from product oriented ownership model to a service oriented access or streaming model. This shift took place in parallel to transformations in both accessing and consuming music facilitated by digitization of music production and distribution. In this model, digital music is offered as a service rather than a product to be purchased or downloaded like in à-la-carte offers of a song or album in the digital retail stores such as iTunes (Dörr, Wagner, Benlian and Hess, 2013). Instead, music files are streamed to the user's device while she is consuming the music on the go. Thus, songs do not have to be permanently saved on the user's device anymore. Rather, music listeners can access to music at unprecedented levels whenever they want by using the libraries located on the servers of online music services. In this way, according to Dörr et al. (2013), the digitization of music distribution has led to the commodification of *access* to the music object rather than the commodification of a music file itself. As a result, the product-based logic of music industry is shifted to access-based service logic.

music industry. Parallel to this, I researched music platforms and familiarized myself with music industry dynamics. The Director of Economics at the pioneer music streaming service, Spotify, supported this research by providing me with a free subscription for Musically - a music industry newsfeed site. Following this website regularly enabled the observation of the music industry at work on a daily basis. This helped in identifying partnerships and innovations in the industry, consequently paving the way for me to reach out to an innovative start-up called Cubic.fm.

Cubic.fm was one of the winners of an annual music start-up contest called MidemLab. Midem is the leading music industry event for the music ecosystem, and it rewards digital innovations in the entertainment field worldwide. MidemLab first mentioned Cubic.fm in news coverage in which it declared the 30 music start-ups that need to be followed in 2014[3]. Here, Cubic.fm was described as a crowd-sourced radio aiming to eliminate the fragmentation among streaming platforms by centralizing data about music consumption: *"Cubic.fm is a social music platform aiming to connect Deezer and Spotify within a single global network. With a unique approach, the service enables users to "cube" (contribute) songs, and listen to activity/genre based radio stations that are algorithmically created from this user generated data. Users build their profiles or get reputation based on their listening and sharing habits."* Cubic.fm then won the Coup de Coeur award in the MidemLab contest for their innovative platform. This music platform appeared as a typical representative of social media, and thus as an eligible case for this research. I met with the co-founders (developers) and was granted access for field study.

iii. Case Study Design

According to Yin (2003), there are several steps in case study design. These are comprised of: 1) definition and design; 2) preparation and data collection; and 3) analysis and conclusion. This process starts with the development of a preliminary theory to guide data collection and analysis. This is followed by the selection of the case. Guiding theory in this research is the preliminary conceptual framework that outlines the platformization of cultural intermediation. The research is built on a single explanatory case study on a typical social media platform. The unit of analysis is determined by purposive sampling, first as the

[3] 'Midemlab Reveals The 30 Music Startups To Watch In 2014'. Available at: http://musically.com/2013/12/17/midemlab-2014-music-startups-finalists/

platform. Then the process is also examined as another unit of analysis. Data is collected and organised through the corpus construction technique; it is collected from multiple sources of evidence. Data analysis started during corpus construction and is conducted using thematic analysis. Yin (2003) advocates using a preliminary theory as the basis for building the explanation and also as a technique to maintain quality. This preliminary theory used for explanation building helps to increase the internal and external validity of the research. Construct validity is increased through different units of analysis as well as different sources data collected as empirical evidence. Table 4.1 below summarises the research design.

Research Design	**This Study**
Time Period	April 2014- December 2015
Sampling Method	Purposive Sampling for unit of analysis Corpus construction for data selection and coding
Data Collection & Data Sources	Semi-structured interviews (26) Screenshots of all different versions of the platform (more than 400) Observation notes Emails Drawings Documents (news coverage(18); blog posts (12); media/press kits (4) Platform's support portal Company's social media accounts
Data Analysis	Corpus Construction Thematic Analysis
Reports	Case Study Narrative (Chapter 5) Findings (Chapter 6) Analysis (Chapter 7) Discussion (Chapter 8)
Quality Criteria	Corroboration meetings held for validation of empirical narrative and thematic analysis Thick description (Chapter 5 and 6) Data triangulation

Table 4.1. Research Design

The case study design is initially structured according to the conceptual framework and planned accordingly. It is then redesigned in accordance with the unexpected findings in the field study. According to the initial plan, the empirical field study was designed to take place over 6 months. During this time, I planned to first explore the context of the company by learning more about its history and the background of the Cubic.fm platform. It was against this background that the researcher was supposed to proceed to the explanatory phase in

which she would collect data to use in answering research questions. As explained in detail in the upcoming section, the exploratory phase has taken more time than was originally planned. This necessitated the adjustment of the initial plans regarding corpus construction. Similarly, corpus design was altered one more time during the explanatory phase. The next section describes this process in detail.

4.4. Data Collection

This section discusses the methods of data collection used in the case study. Despite their limitations for generalizability of findings case studies are recognised to make important contributions for an in-depth study of the complex contemporary phenomenon on such as social media (Merriam, 2009; Yin 2003). According to Patton (1990) an important strength of case studies is enhancing data credibility through multiple data sources. This according to Yin (2003), helps converging data from different sources for triangulation.

Triangulation is as a strategy to test validity of data. It refers to the use of multiple data sources to build a widespread understanding of the studied topic (Denzin 1989; Patton, 1999). Data triangulation in particular refers to the use of different data sources and it is not the same with using different methods to collect data. Denzin (1989) describes the concept of 'data triangulation' as integration of different data sources, distinguished by time, place, and person. In other words, he suggests studying collecting information at different times and places and with different persons. Thus is considered as an effective strategy for researchers to take different perspectives on the phenomenon under investigation.

Thus collected from multiple sources comparison of different types of data enhances data quality through convergence and the confirmation of findings (Knafl & Breitmayer, 1989). In this regard, it is used as a strategy for crosschecking information and conclusions using different types of data complementary to each other. This sheds light on different aspects of the studied topic. As such, Johnson (1997) argues, when different procedure or sources are in agreement researchers achieve "corroboration."

Yin (2003) suggests using six prominent sources to collect data. These are: documentation, archival records, interviews, direct observations, participant observation, and physical artefacts. Data collected from these sources are converged in the analysis rather than being analysed independently. Each data source, according to Baxter and Jack (2008) constitutes a piece of the "puzzle," and each piece contributes to building a better understanding of the

investigated phenomenon. This convergence adds strength to the findings as the various elements are interwoven together to create a holistic understanding of the case.

4.4.1. Using Multiple Sources of Evidence

In light of this, 26 semi-structured qualitative interviews were conducted with the 5 people. As discussed in the next section, these are the people who have the in-depth knowledge of the object of the study, that is the platform system of the chosen company. All these informants were involved in the development of the product-platform of the chosen company. They are the experts who could explain the inner workings of the chosen platform in relation to how it is configured. Explaining this configuration also requires giving information about the technical and business concerns encountered during the development. Interviewing people, who have expert knowledge of the studied topic, is a recognized way to collect such information (Flick, 2009; Meuser and Nagel 2002; Bogner and Menz 2002). This especially holds true if the object of the study is a complex contemporary phenomenon (Yin, 2003) like the sociotechnical design of a social media platform.

Nevertheless, the researcher acknowledges the limitation of interviewing 5 people for the generalizability of the findings. Despite being conducted in limited numbers, interviews can still be a useful way of data collection if researchers reflect on these limitations. In order to overcome this limitation the strategy of triangulation of data sources and data types is used with the aim of increasing data credibility and the 'truth value' (Baxter and Jack 2008). Triangulation helped balancing out the limited number of interviews. Moreover, the same people were interviewed several times at different points during the field study following major events in the evolution of the platform system. Thus, in line Denzin's (1989) views, data sources are also differentiated by time and rapport with informants was established and resulted in collecting multiple perspectives (Krefting, 1991).

Consequently, data triangulation and having a prolonged exposure to the phenomenon under study helped to overcome the limitation of interviewing limited number of people. This allowed collecting detailed information regarding the issues that cannot be understood otherwise (i.e. how developers tweak the platform system in response to users reactions and other external factors, this, how did they use data-based means and reporting tools such analytics to understand user behaviours and how this is implicated in the optimisation of platform system).

In-depth interviewing is complemented by observations. The researcher held brief observation sessions following most of the interviews. Those helped the researcher observe different aspects of the phenomenon of interest to draw information, which was not obtainable from other sources. However, this method is not also free from limitations. The main problem is that, not all phenomena can be observed in situations. This is especially true if the studied organisation is building a software product that requires coding activities of developers. Comprehensive knowledge processes in such environments are difficult to observe and events or practices may seldom occur (Flick, 2009). To address these problems, additional interviews are integrated to the observation sessions. This allowed exposing the necessary information shedding light on the background of observed events.

In addition, the fieldwork included the collection and analysis of documentary sources. This included relevant, non-confidential publically available documents that were available for this research (i.e. company documents, online documents about the company and the platform, emails etc.). Documents are more concrete than interviews or participant observation in that it is usually clear what has been said, and by whom. But, even texts can be ambiguous (Myers, 2004). Moreover, it is argued that documents should be seen as "communicative devices produced, used, and reused for specific practical purposes, rather than as 'unobtrusive' data in the sense of bias-free data" (Flick 2009, p.262). Nevertheless, they can provide a valuable addition to other forms of data, if the contexts within which they are produced and used is taken into account. Visual data was another form of evidence used in complementary to other data sources. More specifically, publically available videos about the company are collected to help data triangulation as visual data are recognized as valuable complementary sources for qualitative research (Rosenstein, 2002; Flick, 2009). Visual material like photos and videos are used with increasing frequency in complementary to spoken word and the textual medium (Flick, 2004). Last but not least, company's product-platform (Cubic.fm website) constituted another data source for data collection (Marshall and Rossman, 1995; Yin 2003). Indeed, digital artifacts like this are used as a valuable data source used in contemporary case studies, especially if it is a crucial part of the studied phenomenon (i.e. Levenshus, 2010) or if the digital artifact itself is the very phenomenon under investigation (i.e. see Alaimo, 2014). Table 4.2 below provides a summary of the data sources, and the section following it explains each of these in more detail.

Data source	Details
Interviews	26 Semi-structured and unstructured interviews with 5 members of Cubic.fm team. 4 Interviews with CEO and Developer Onur Y. 4 interviews with the Second Developer Ozgur A. 4 Interviews with the Analyst Baris A. 7 Interviews with the Analyst Erdem G. 1 Interview with the Mobile App Developer Rifat O. 2 joint Interviews with the Analysts Baris A. and Erdem G. 1 Group meeting with the whole team (See Appendix A)
Observation	Notes taken during observation sessions in the site visits following the interviews. This included generation and Analysis of Analytics Reports (i.e. cohort analysis) showing the changes in weekly user engagement (i.e. percentages showing the weekly changes in the number of users acquired, retained, churned etc.
Artefact	Notes on 1) the researcher's experience as a user with different versions of the platform, 2) observation of users' interaction with the platform as well as the interaction amongst users. 3) Notes taken to give feedback Cubic.fm team (as a form of user feedback) Screenshots of: • Features of different versions of the platform • Output of different data inputs and search queries • Email notifications and pop-up notifications about the users' reactions to the researcher's playlists
Emails	108 email exchanges on 20 different topics
Videos	11 Video Content publicly available on YouTube and Vimeo (See Appendix B)
Documents	Company documents • Infodeck/presskit • Sample database queries • Documents illustrating data aggregation procedures Other sources: • Product Hunt Launch Q &A with developer community • Social Media Pages • Company's blog • Platform's Support Centre • News Articles & Blog posts (See Appendix C)

Table 4.2. Empirical Data Sources

4.4.2. Interviews

Initial contact was made with the CEO of the company to assess the suitability of the company and its product-platform for the research. This meeting also involved a discussion of the availability of the team to accommodate this study in parallel to their development and marketing efforts. A second meeting with all five members of the team followed this initial one where the researcher explained the purpose of the research, discussed the research procedure, consent of the participants is received and a timetable was created to arrange weekly meetings. Thus, the researcher agreed with the team to conduct interviews with each member of the team on site, to observe the interactions in their office settings before and after the following the interviews and by participating in some of their meetings. Interviews were typically 45 minutes to 1 hour in length and they are recorded using a voice-recorded device. These are conducted in on site in the company's office. The researcher transcribed all interviews with the help of online free transcription software before they were coded manually. 26 interviews with 5 informants were conducted. These include one meeting with the whole team and two joint interviews with two analysts.

This thesis explores social media's sociotechnical configuration by focusing on a representative platform's design rationale and data operations. Finding answers to the research question posed in this study requires having a detailed account of the perceptions, expectations, and interests underpinning the development of a social media platform. These can be best explored by talking to the key actors involved in the design and development of the studied platform-product. Interviews were chosen as the main method of data collection for this study. This is a widely used method in qualitative research. It is expected that, the informants are more likely to express their viewpoints in an openly designed interviews than in a standardized questionnaire (Gubrium and Holstein 2001; Flick, 2009).

In this regard, interviews were conducted with the 5 key people involved in development of Cubic.fm platform. This included four co-founders, namely: two developers (the CEO of the company being one of them); two analysts who were PR and marketing specialists responsible for user engagement and growth; and one mobile app developer. These Informants provided different viewpoints regarding the same topic based on their background, responsibility scope of their work, their role in the company etc. Beeke (1995 cited in Flick, 2009) describes experts as people "who are particularly competent as authorities on a certain matter of facts" (p. 8). According to Bogner and Menz (2002) experts can also be staff members of a particular organization under scrutiny who have a specific

function and a particular professional experience and knowledge. It is considered that these people have "technical process oriented and interpretive knowledge referring to their specific professional sphere of activity" (ibid. p.46).

For instance developers were trained as computer scientists and their understanding of the system was very different than that of the analysts, who had a background in management and economics. This was also evident in the distribution of their roles. While the analysts focused on user research and experience their insights informed how developers worked to improve system functionality or design and deploy new features by coding. Therefore the insight provided by developers and analysts were completely different than each other. This reflects how Bogner and Menz (2002, p. 46) describes expert knowledge not only consist of specialist knowledge, but has the character of practical knowledge.

As emphasized by Meuser and Nagel (2002), when interviewing experts, interviewees are of less interest as a person than their capacities as experts for a certain field of activity. Only these actors could provide the information about different technical and business related concerns that shaped the specific design logic adopted when building this particular social media platform, as they are people who have a complex stock of knowledge about the topic under study. This knowledge includes their perceptions and assumptions that are explicit and immediate and which interviewees can express spontaneously in answering an open question (Flick 2009).

Questions to Ask

It is noteworthy to describe here how developments during the field study shaped the data collection and analysis. Data collected though the initial set of semi-structured interview questions the interview guide. Following the interviews data was transcribed and analysed to generate follow up questions. As data are collected and analysed, the researchers' interpretations of the data are occasionally shared with the participants, and the participants have the opportunity to discuss and clarify the interpretation, and contribute new or additional perspectives on the issue under study (Baxter and Jack 2008). At points this process led to generating a completely new set of questions that has given new directions for the upcoming interviews. These were the times when the researcher decided to focus on a new topic brought into her attention by the informants during the previous interviews.

This is one of the recognized problems in conducting interviews. Merton and Kendall (1946) suggest for conducting the interview incorporate some targets such as specificity and depth

versus range. But they recognize that they cannot be matched in every situation as fulfilling these criteria cannot be realized in advance. How far they are really met in an actual interview, according to Flick (2009: p.154) "depends to a great extent on the actual interview situation and how it goes off. These criteria highlight the decisions that the interviewers have to make and the necessary priorities they have to establish ad hoc in the interview situation. They also mention there is no 'right' behaviour for the interviewer in the focused (or any other semi-structured) interview."

The evolution of the platform and significantly different forms it took during the study was not foreseen neither by the researcher nor the informants. Changes in features were expected to a certain degree and the possibility of pivoting ideas was acknowledged beforehand. Nevertheless because the decision to release of new versions were made following the lean start philosophy that advocates failing earlier, the company was able to take decisions in sudden and agile ways based on the user feedback and their growth reports. In this regard, interviews in every few months involved a new and unexpected change about which the researcher were informed on the site where she went to collect more data about the topics discussed in the previous interview based on a new set of questions. In such cases the researcher used open-ended questions to find out the reasons behind the decision to release a new version and/or the changes that were planned previously but then abandoned in between the consecutive interviews. In doing so, 'retrospective inspection' (Flick, 2009) helped the researcher to increase specificity and collect more detailed information by asking informants to describe changes as a chain of events.

This reflects that how preliminary analysis actually begins during data collection where researchers need to make decisions about narrow down the their cope or, develop new set of questions and/or re-evaluate their data-collection instruments (Bogdan & Biklen, 1992). It also illustrates advantages of conducting case study to investigate a contemporary phenomenon like social media: enabling the researcher to gain an complete view of a certain phenomenon or series of events (Gummesson1991; Yin, 2003) by capturing the emergent and immanent properties of the issues in organisations, the studied topic, and/or the flow of activity, where it is changing very fast (Hartley, 1994).

Like all methods of qualitative data collection, interviews involve a degree of subjectivism in their interpretation (Holstein & Gubrium, 1997). Gaskell (2000) points out the limitations of interviews as a data collection method, which amount to the fact that all information is gained

through the eyes of the interviewees, so it may be limited or biased. Even the process of transcribing interview recordings can be seen as an act of construction and sense making (Hammersley, 2010). Thus, the interviewer plays a key role in constructing meaning as well, for example by choosing specific questions, or by writing up the interview findings in a certain way (Kvale & Brinkmann (2008). Interviews can still be a useful way of data collection if researchers reflect on these limitations. In this regard, the subjective nature of the interviews is acknowledged in this study. It has been pointed out, however, that interviews are in no way objective and should not be seen as a method to reveal an inherent truth hidden in a case to be studied.

4.4.3. Observations

Beside the interviews, brief observations sessions were conducted following the interviewees in their office. Observation sessions needed to be complemented by other data sources due to the difficulty of observing tasks such as coding. Moreover, significant events seldom occurred in a software development company like this. As a way of solving these problems, additional interviews of participants are integrated into the research program, which allowed exposing stocks of knowledge that are the background of observable practices (Flick, 2009). Thus the researcher focused on observing the tasks team members were working on (i.e. use of analytics tools, generating reports for weekly meetings, analyzing user feedback etc.). Despite occurring rarely, these sessions also included observation of discussions between team members about other platforms or about the development of a particular feature. Consequently, data collected through this method is to a large extended grounded in participants' verbal statements about certain relations and facts rather than observation of actions. Nevertheless, these sessions allowed the researcher gather valuable information complementing the information provided by the informants about their daily activities and how these are implicated in optimizing the platform system.

This included observing how they different forms of user feedback are retrieved from different channels (i.e. social media, support center, emails, surveys etc.) and analysed, how analysts tracked user behaviours and analysed behavioral data using analytics tools to generate reports. These observations sessions allowed asking informants follow-up questions regarding how a particular report inform their decisions such as how to improve user experience by changing a system functionality and what this would mean for the business interests of the company etc. Some follow up questions could be only asked during these

observation sessions as the researcher come across new information which was not provided by the participants during the interviews. Informants' answers and comments were noted by the researcher as necessary and used complementary to interview material. Thus, observations helped the researcher to collect data about different aspects of the phenomenon of interest, which was not obtainable from other sources.

4.4.4. The Artifact

Artifacts are recognized as one of the multiple sources of evidence (Yin 2003). An important source to gather evidence in this sense was the object of the study itself, that is the sociotechnical platform of the company. The researcher was given access to different versions of the platform as a user. This included the private beta versions before their public launch as well as the publically available website of the platform. Thus I had the opportunity to observe the ways in which platform works at the front end by: a) drawing on my own experience with the platform; b) by interacting with other users and by observing the interactions amongst them. This involved attempts to observe how users react to metrics (i.e. observing the activities of users with large number of playlists and followers.). When using and observing the website I made notes and took screenshots for later review to be used once themes in the interviews are identified. As a result, a large number of screenshots were taken to be able to make a longitudinal observation of users' interaction with the platform. The Web site observations and themes were categorized in a separate spreadsheet and used for crosschecking conclusions drawn from other sources.

4.4.5. Documents

Documents constituted another important data source. Like other approaches in qualitative research, documents are also used complementary to other data collection methods like interviews (Flick 2009). Collected documents include relevant and non-confidential documents that were made available for this research; publically available documents collected from online sources (Alasuutari et al., 2008); and emails (Flick, 2004) exchanged with the informants throughout the fieldwork.

Documents can be found in the form of printed text or in the form of an electronic file (Flick 2009). Internal company documents made available to the researcher included both types of texts. This included press kits, /drawings and sketches of the database and data flow diagrams, example SQL queries, documents explaining data aggregation criteria, documents about the previous product of the platform, sample reports etc. Other documents were

collected from online sources hence they were all in digital format. These sources included the company's blog, platform's support centre, social media pages of the company and other online materials such as news articles, blogger posts, interviews etc. Lastly, the researcher used email exchanges with the informants in complementary to other data, as emails are recognised as new forms of communicative tools and hence valuable data sources (Flick, 2004). They helped the researcher in clarification of points discussed in the interviews, following updates about the upcoming changes and the dates, receiving additional material from the informants, which they thought, would be of help for the study and so on.

A crucial point about using documents as a data source is being aware of the broader context within which they are produced and used. According to Flick (2009), "*Documents are not just a simple representation of facts or reality. Someone (or an institution) produces them for some (practical) purpose and for some form of use (which also includes a definition of who is meant to have access to them). When you decide to use documents in your study, you should always see them as a means for communication. You also should ask yourself: Who has produced this document, for which purpose, and for whom? What were the personal or institutional intentions to produce and store this document or this kind of document*?" (Flick, 2009: p. 257). From this perspective, different types of textual material helped the researcher to examine how different narratives about the company and its product platform were created throughout the evolution of the company. Thus the researcher avoided focusing only on the contents of documents without taking their context, use, and function into account. Rather they are seen as a form of communication with different stakeholders –existing users, potential users, potential collaborators, investors, press, labels and brands (for affiliation agreements etc.). Thus, collected evidence involve different messages given to these external actors as the platform evolved from a recommendation based discovery platform focused on activity/mood to a fully functioning social media platform reliant on users and finally to a bookmarking tool used for building personal library.

4.4.6. Videos

Apart from data collection oriented to the spoken word and the textual medium, and beyond observation in the field, new media are becoming increasingly important as data (Flick, 2004; Rosenstein, 2002). Photographs, films, and videos are increasingly used as important forms and sources of data (Flick, 2009). In this sense, video content was another form of data collected for the analysis during the study (See Appendix A for the details and the URL links

of these videos). This included publically available videos on the social media platforms YouTube.

Some of these videos were generated by the Cubic.fm team and made available to public. These were mainly in the form of demonstration videos called 'demo's. There are separate videos for different versions of the platform explaining its value propositions for the users. Moreover, Cubic.fm participated in different start-up competitions throughout the evolution of its product-platform. YouTube features video content about the company pitches in start-up competitions with different versions of the product-platform. In these videos, one of the team members explains their product to the start-up community and investors, and answers their questions. These videos involve information about the overall design logic of the platform in relation to the company's business model, value proposition, and growth strategy. Finally, there are videos that involve interviews with the team members conducted at different stages of the evolution of the company's product-platform.

It is noteworthy to mention that, similar to the documents, visual data such as video content are also recorded in different contexts and/or with different audience in mind. Flick (2009) argues that they can be very instructive addition to interviews or observations as long as the researcher can conceptualize the relations between explicit content and the implicit meaning, and how to take this into account in the interpretation of data. From this viewpoint, publically available video material helped the researcher to gather information about how the product-platform is presented to different set of audiences such as users, investors, and other actors in the industry. All these stakeholders have different interests and these videos represent how the company communicate with these stakeholders to address their needs and concerns. In this regard, this audio-visual material provided important complement information to triangulate data collected from other sources.

4.5. Corpus Construction

Corpus construction is a data collection and selection technique. It originated as a linguistic technique and was adopted in social science by Bauer and Aarts (2000) to investigate *unknown* subject matter through *known* categories. In this technique, external categories are used to gradually identify and outline the core themes of the phenomenon under investigation. This process is based on a continuous feedback between theoretical constructs, empirical data, and the on-going interpretation of research findings. Thus data collection is

rendered a dynamic process, gradually shaped along the way in light of new findings.

4.5.1. Sampling

Corpus construction is another term for what is referred to as a purposive or theoretical sampling method (Corbin and Strauss, 2008). However, in contrast to theoretical sampling, corpus construction is based on a two-dimensional investigation in which the researcher aims to account for a variety of unknown phenomenon through known categories. These terms are basically used to emphasise how this method differs from random sampling.

Random sampling depends on a previously set sampling arrangement for a given population whereby the population is built based on the distribution of previously defined variables. Contrary to this method, purposive sampling relies on initial research findings to inform the subsequent data collection and selection strategy. Thus it is based on an iterative process in which an initial set of data shapes the criteria for further data selection and allows the gradual collection and selection of sample data for analysis. Such an approach enables further exploration. When, for example, it is not easy or insightful to construct a sample for a given population based on predefined variables, the research can progress through an initial exploration of the empirical domain – i.e. the configuration of social media platforms. Thus, a corpus is designed and constructed through a cyclical process (Bauer and Aarts, 2000; Biber, 1993) in which the data and findings iteratively inform the design while the research results take shape gradually.

In this regard, the rationale behind the data collection and selection process should take into account the idiosyncrasies of the unknown phenomenon under scrutiny. The focus should be on *variety* in data that can help to explain different aspects of a phenomenon, rather than its distribution. Thus a corpus, composed of different kinds of data (e.g., text, still or moving images, sound, etc.), can represent the regularities as well as irregularities in a population and open up a new space for further exploration. This connects to the issue of *representativeness.*

4.5.2. Representativeness

Representativeness refers to how quality is achieved and defined as "the extent to which a sample includes the full range of variability in a population" (Biber, 1993, p.243). In order to ensure representativeness in corpus construction, Bauer and Aarts (2000) suggest proceeding stepwise by: a) selecting, b) analysing, and c) selecting again in order to "maximize the

unknown variety of representations by extending the range of the known variety until saturation is achieved." This suggests that qualitative researchers deal with the unexpected and uncertain through a continuous process of selecting and interpreting data. This helps to map the investigated phenomenon gradually in a systematic manner. In this process, previously identified categories guide the researcher as structuring devices. These are applied to the unknown to focus attention on particular aspects of the phenomenon at hand while navigating through an unknown territory. In this way, these external categories serve as the backdrop against which the researcher becomes able to single out unexpected or surprising occurrences. Thus the corpus design can be updated to collect more data on those aspects if necessary. In the following, I discuss the rules and criteria used to identify the external categories used in corpus construction and explain the evolution of the corpus design.

4.5.3. Corpus Design

Previously known categories are used in corpus construction to segment the population. These external categories help to overcome the dilemma of constructing a representative sample for an unknown phenomenon. Theoretical accounts that informed the framework devised for this study also shaped the early corpus design. Thus the corpus is initially constructed based on two core categories in the exploratory phase: 1) the overall design logic and 2) the backend data-based operations underpinning a social media platform's matchmaking function.

The exploratory phase was originally planned to last around a month. However, once this phase started, the researcher realised that the platform was in a transitory phase, switching from an 'activity-genre-based crowd sourced radio' that connected two prominent streaming platforms (Spotify and Deezer) to a more generic social media platform (in the form of a social music discovery platform), which aimed to connect as many music services as possible. In this first round of data collection, the researcher simultaneously attended to data coding in order to gain a deeper understanding of the reasons and the rationale behind this change. This initial process of data elicitation and coding revealed that this decision to change platform configuration was mainly a move towards turning the platform into a *self-sustaining* 'social media platform' that is *completely reliant on* user-generated data and content. Because it was conducted in this transition stage, new questions emerged to explore the logic behind this change. The exploratory phase was consequently extended to take four months. This period resulted in a new design for corpus construction. The researcher

decided to collect and select data based on two new categories: 1) the *platform* and 2) the *process*. *The platform* entailed the platform's design logic and the data operations that sustain it, whereas *the process* delineated the changes made in the platform design as well as the reasons and rationale behind them. Thus the initial exploratory phase served to contextualize the research and adjust the corpus design in accordance with the preliminary findings.

The explanatory phase gave rise to another change in corpus design. After focusing data collection on the product and the process, one more dimension appeared as relevant for corpus construction. This new dimension was identified following another core change made in the platform design. Such a change at first puzzled the researcher, since the research was initially designed with the expectation to study a stable platform system. The initial idea was to understand the relationship between the frontend and backend of social media. However, new developments resulted in new findings. Perhaps one of the most prominent developments was the recognition that neither the frontend nor the backend of a social media platform are fixed. Thus the corpus was reshaped one more time. In addition to the previous two dimensions (the *platform* and the *process*), a new dimension was identified as the different versions/forms the platform has taken over time. Thus a *time- & form*-based dimension was added to the design. Table 4.2 below summarises the evolution of the corpus design.

Corpus Design	**External Categories**		
Initial Phase	1. Design (frontend) 2. Data operations (backend)		
Second Phase	A. Platform	1. Design 2. Data operations	
	B. Process	*i. Evolution of the Platform*	
Final Phase	A. Platform	a) 1st Vr Cubic.fm	1. Design 2. Data operations
		b) 2nd Vr Cubic.fm	1. Design 2. Data operations
		c) 3rd Vr Cubic.fm	1. Design 2. Data operations
	B. Process	*ii. Evolution of 1st Vr*	▪ RBC
		iii. Evolution of 2nd Vr	▪ RBC
		iv. Evolution of 3rd Vr	▪ RBC

Table 4.3 Summary of the Evolution of the External Categories used in Corpus Design (Vr

Stands for *Version;* RBC stands for *Reasons behind Changes in Configuration*)

In addition to these external categories, the implementation of corpus constructions also followed the three criteria identified by Bauer and Aarts (2000): 1) relevance, 2) homogeneity and 3) synchronicity. These are used in building sub-corpora depicted in Table 4.2 above through different types of data, ranging from observation notes, interviews, documents, and other audio-visual material available about the platform.

Relevance criteria basically helped in distinguishing between what is relevant and irrelevant data based on the research questions and the theoretical lenses guided this study. However, as previously described, initially chosen categories evolved over time as unexpected developments in the field, and some surprising findings paved the way for adding new dimensions to the corpus design. In this regard, the criteria for relevance in data collection were continuously updated. Accordingly, the question regarding 'what is relevant' and 'what is not' also changed over time, as different new sub-corpora emerged over time and helped to explain different aspects of the phenomenon. As a result, the unit of analysis evolved from the platform *per se* (the logical design and data operations) to include *both* the *platform* and the *process* in which it is developed and optimised. Thus, the corpus evolved over time.

Homogeneity refers to the construction of a set of homogenous corpora that are designed to organise the data elicitation by maximising the variety within the case. This refers to maintaining each corpus as homogeneous by building different corpora for each data (i.e. documents, images etc.). In the analysis, I mainly relied on interviews and applied this principle to the other data gathered from various different sources. I used these additional types of data for triangulation. This helped to verify and increase the explanatory power of the corpus.

The last principle proposed by Bauer and Aarts (2000) is *synchronicity*. This refers to *time homogeneity*. This principle was not a very relevant criterion at the beginning of the research. However, it emerged as a significant factor over time and used in adjusting corpus design. Thus it informed the identification of the new external categories for building additional sub-corpora based on the unexpected findings. It came as the biggest surprise to observe the reconfiguration of the platform system three times, with several other minor changes accompanying each configuration. Perhaps one of the most interesting findings of the research thus emerged as the dynamic and unstable nature of social media platforms, which points to a need for further research on the reasons and drivers behind their dynamic

structure.

The first version of the platform was, in a way, the launch of the platform in the form of a minimum viable product (MVP).[4] The aim behind this was to understand the needs and concerns of users and the industry, to test platform functionalities in real-time, and to address those needs and concerns in the best possible way. This approach to optimising platform configuration through real-time feedback-based hypothesis testing continued throughout the field study. Toward the end, it resulted in the redefinition of the platform one more time as a personal music library, and the corpus design evolved accordingly. Consequently, time-form based changes made in platform configuration have also been taken into account to inform the corpus design in the explanatory phase. Table 4.3 below summarises the evolution of corpus construction based on the synchronicity principle.

Corpus	**Time period**	**Unit of Analysis**	**Output**
Initial Design	April- October 2014: 1 month exploratory phase followed by 5 months long explanatory phase	Platform (and context)	Corpus
First Round	April-August 2014: Actual Exploratory Phase that lasted 4 months	Platform (and context)	Adjustment of unit analysis and design of corpus construction as: 1) Platform 2) Process
Second Round	September 2014-March 2015	Platform and Process	Adjustment of unit analysis and design of corpus construction as: 1) Time-form based changes in platform 2) Factors that gave rise to these changes in the development process
Third Round	March 2015-December 15	1) Platform 2) Process	The final form of corpus

[4] A minimum viable product (MVP) is "a product with just enough features to satisfy early customers, and to provide feedback for future product development." Definition is retrieved from Wikipedia, available at: https://en.wikipedia.org/wiki/Minimum_viable_product

		Based in three different time periods marked by two fundamental changes made in platform configuration	

Table 4.4 Evolution of corpus construction based on the synchronicity principle

4.6. Analysis

Previous sections discussed the fundamentals of research design. They outlined the epistemological and ontological stance of the study, discussed methodology, and introduced the chosen case. This was followed by an explanation of corpus construction. In this section, I discuss the analysis of empirical evidence collected through the chosen case study.

In the analysis process, I mainly relied on the analytical abstraction method outlined by Miles & Huberman (1994). According to Miles & Huberman, there are three flows of activity that make up the process of qualitative analysis based on analytical abstraction: i) data reduction (summarizing and packaging the data), ii) data display (repackaging and aggregating the data), and iii) conclusion drawing and verification (developing and testing propositions to construct an explanatory framework). Analysis of the case study evolved gradually following these stages of analytical abstraction.

4.6.1. Data Reduction

In the first stage of analysis, the corpus is prepared for analysis. It is noteworthy to mention that this process started during data collection. Thus corpus construction and analysis continuously fed each other. From the initial weeks onwards in the data collection and selection process, the researcher transcribed recorded interviews and summarized the main points of each interview. This helped in the construction of an initial coding scheme, which guided the description of the preliminary case findings. Thus the initial codebook was developed. This preliminary version of the coding process was mainly built on the conceptual framework. It focused on the events concerned with the design logic and underpinning data operations reported by the interviews. After observing in findings new events that could not be explained by the initial conceptual lenses and the accordingly chosen categories, both the corpus design and accompanying coding scheme were extended and adjusted. Thus the interviews are transcribed and coded manually on a continual basis. This stage served as the

basis for the case study narrative upon which the subsequent analysis and discussion are built.

4.6.2. Data Display

In the second stage, the analysis is concerned with data display that is repackaging and aggregating the data (Miles and Huberman, 1994). At this stage, first the themes and trends emerging from raw data are identified. This is followed by the development of new conceptual codes or the refinement of existing ones by merging the summaries of codes in a second round of coding. After this stage, the initial bulk of mainly descriptive codes was cleaned, merged and reduced using the criteria identified for relevance (research question and corpus design). This helped in repackaging data for the next level of conceptual abstraction, while paving the way for identifying additional findings on the way. Thus a higher level of analytical abstraction was achieved through the ideas developed in memos, personal notes and mind maps and continuously adjusted by synthesizing theory and data on an ongoing basis. Consequently, this second round of coding resulted in the identification of the major findings of the study, as presented in Chapter 6. This moved the analysis from a descriptive to a more conceptual level, generating two major findings: 1) the challenges and concerns faced during the development and optimization of a social media platform and 2) the tools and techniques used to address these challenges.

These relatively more conceptual themes were identified through thematic analysis (Boyatzis, 1998), based on a cyclical process that was continuously informed by both theory and data. Thus it served as the basis for the subsequent analysis and discussion. As a qualitative method, thematic analysis is built on an inductive rationale and employs an open coding technique. Themes are gradually devised from data based on trends and patterns emerging as distinct data clusters. Thus, in addition to the corpus construction technique, thematic analysis is used as an organizing device to describe data through the interpretative filter of the researcher (Braun and Clarke, 2006). The final version of the coding scheme is thereby built on the active interpretation of data in light of the initial theoretical considerations and the emerging themes.

The researcher was prepared for and hence expecting to find some unexpected findings at the beginning of the study due to the dilemma intrinsic to qualitative research, arising from researching an unknown phenomenon. However, there were more issues involved in the process of developing a social media platform. Developers reported that social media

platforms, usually built through lean start-ups, typically launch a minimum viable product to get real-time feedback for ongoing product development and product optimization. The study revealed from the very early stages onwards that optimization is a crucial part of the process in which a social media platform is developed. Indeed, it never ends. These platforms are continuously developed and adjusted based on the perceived user needs and emerging business requirements and opportunities of the broader digital ecosystem within which they are embedded. In this light, the ideas outlined in the preliminary framework are further developed through the identified themes based on the more conceptual round of coding. In this final round of coding, the relationship among the codes and conceptual themes becomes more prominent. Thus the resultant explanatory framework is developed based on both the expected as well as the unexpected findings.

Data display plays an important role in this gradual process of analytical abstraction. The first round of data re-packing underpinned the descriptive display of the empirical data in Chapter 5. Here, vignettes from interviews and the screenshots from different versions of the studied platform are presented within the case study narrative. These are accompanied by images of some of the company document, press coverage about the platform, links of video content that involves the presentation of the platform and its business model to investors, or sketches of the database and data processing. This is followed by a more thematic display of data in Chapter 6, in which additional findings were presented. Thus the analysis evolved into a more conceptual level.

4.6.3. Conclusion Drawing

In the final stage of analysis, an explanatory framework was constructed based on the preliminary framework and the themes that emerged from analysis of the findings of the case study. This is achieved through: 1) reducing the bulk of the data (observable events) for the analysis of patterns in it, then 2) summarizing findings to identify key themes, and finally, 3) testing assumptions regarding the mechanisms producing the observed events. Thematic findings were discussed with the informants in the case study twice. These final meetings served for the refinement and validation of the conceptual ideas as the mechanisms generating the observed events in the case study. Thus the explanatory framework delineating the deep structure underlying the platformization of cultural intermediation took its final shape gradually through corroboration.

As discussed previously, the retroduction principle guided the process in which these mechanisms were identified and corroborated. This process started with the descriptive answers informants gave to the initial questions of the researcher. These answers were based on the interpretive judgments of the participants; hence they required an analytical repackaging. This is followed by the researcher's theoretical interpretation of the descriptive accounts. Thus drawing on preliminary concepts, a theoretical account of the situation was developed. Then the researcher elaborated on the relationship between the concepts and themes identified as key components constituting the mechanisms. Thus alternative explanations were tested. Finally, the explanatory power of the final framework was validated by discussing with informants how the identified mechanisms manifest themselves in the case study.

4.7. Conclusion

In this chapter, I discussed the research strategy this study adopts to answer the research questions. I first outlined the epistemological and ontological stance of the study as a combination of constructivism and critical realism. Here I emphasised that this study aims to go beyond the socially constructed aspects of social media platforms and focuses on the construction of kinds. In this regard, the study focuses on revealing the *particular ways* in which social media platforms transform cultural intermediation. This requires deconstructing the sociotechnical configuration of social media platforms, which also entails revealing the assumptions underpinning the overall design logic and underpinning data operations. These are, however, not easily observable phenomena, and hence require going beyond the observable events and scrutinising the mechanisms producing them. In this regard, this study aims to investigate and explain how the events observed in the chosen case study are produced.

Following this section, I explained why case study research and qualitative methods are chosen to investigate the sociotechnical configuration of social media platforms. I explained here that these approaches are adopted due to their flexibility and the opportunity they give to researchers to study complex phenomena like social media in their real life context by drawing on multiple sources of evidence. I then introduced the case study and proceeded to a discussion of corpus construction as the chosen data collection and organization technique. Here I introduced the rules and criteria used in the corpus design and depicted how these evolved over time, as informed by theory and the interpretation of data on a continual basis.

Finally, I concluded the chapter by outlining the data analysis phase. Here I organized the section around the three key stages that guided the analysis: data reduction, data display and conclusion drawing. The next chapter presents a descriptive account of the case study, as conducted based on the research design principles delineated in this chapter.

5. Empirical Case Study

5.1. Introduction

This chapter sets the context of the empirical case study with a descriptive account of the company investigated and its evolving product-platform. The main objective is to build a flowing narrative of the organization, describing its aims, concepts and visions. The narrative constitutes the background against which the major themes of interest will be analyzed. It is worth noting that this is a study conducted while the investigated company was still in its start-up phase, during a critical time frame in which its product-platform was redefined three times. As such, it includes observations of the ways in which the company addressed issues with and concerns about the social platform it strived to build. Because each of the three stages of the platform are markedly different, the case study narrative engages with them as three distinct versions, each one of them dealing with different facets and problems that the company faced.

The chapter is comprised of three parts. Part one provides the *context* for the analysis, giving an overview of the company as a data-driven organization as well as a brief description of its background, vision and business model. The second part of the chapter entails a description of the main object of this research; that is: the design, development and evolution of the company's product-platform. For clarification, this section is divided into two sub-sections: the *platform design* and the *development process*. This section first looks at the three forms that the design of the platform has taken throughout its evolution in light of the changes that rendered them distinct from one another. In the second half of this section, the evolution of the platform is discussed with regard to the challenges the company faced in the process of developing a social media platform in the domain of music.

The second half of this chapter as a whole provides a detailed account of the three versions of the platform, the features and functionalities that distinguish them, and the rationale underlying changes made in the platform system. These observations will be used throughout subsequent chapters as the basis for analysis and as empirical reference points to ground the study's findings and main discussion themes.

5.2. Company and the Context

5.2.1. Overview

Cubic.fm is an Istanbul-based data-driven company founded in 2013 by four young entrepreneurs who studied in the same university. The core product of the company is a global music platform operating free of licensing and content delivery network costs. In 2014, the company was one of the winners at the Midem Conference, an annual event dedicated to the music industry in which 30 finalist start-ups from Europe pitch their products (see Figure 5.1 and 5.2 for press coverage of Cubic.fm as MidemLab finalists and the winner of the Coup de Coeur award). The team also won several other innovative start-up awards from national competitions in Turkey like *Startup Istanbul 2014* and *Startup Turkey Challenge* in 2015 (see Figure 5.3 below)

Figure 5.1 A news coverage of Cubic.fm appeared on the online website of the Magger, a famous cultural magazine in Turkey. Available at: https://www.themagger.com/ajanda/

Figure 5.2. Public announcement of Cubic.fm as the winner of MidemLab. 'Vivendi Awards Its Midem "Coup De Coeur" To Cubic.Fm'. Available at: https://www.vivendi.com/en/press/press-releases/vivendi-awards-its-midem-coup-de-coeur-to-cubic-fm/

Figure 5.3 Cubic.fm team. Retrieved from Startup Turkey Startup Challenge 2015. Available at: http://blog.startupturkey.com/2015/02/startup-challenge-2015-the-most-successful-15-startups/

This company's product-platform, also named Cubic.fm, was first launched as an application in the app studio of the streaming service, Deezer, in the summer of 2013. It was subsequently redefined three times while the company was still in its start-up phase, and several notable changes were made with each version in terms of the product-platform's features and functionalities. Following its launch as a crowd-sourced radio application in Deezer, the same app was launched in Spotify. This paved the way for releasing (1) the stand-alone *crowd-sourced radio platform*, which was eventually redefined as (2) a *social music discovery platform* in the summer of 2014. This version also went through several changes, and eventually the platform re-launched in a private beta as (3) an *online music library for personal use*.

5.2.2. Background

Cubic.fm was not the first product-platform developed by its co-founders. The team previously developed similar platforms in the domain of music recommendation and social discovery. The product they developed before Cubic.fm was a platform called *Tasteer*. Launched in 2012, *Tasteer* was a social recommendation platform for books, movies, and music. After experiencing some problems with this product, the developers decided to focus solely on the music domain. They aimed to increase the quality of recommendations by understanding the specific context of the music listening experience. Tasteer then converted into an online music service that offered personalized genre stations and promoted music discovery via a socially-influenced recommendation system: the 'taste-engine'. However, licensing issues in the music domain did not allow them to stream music content in the way they initially planned.

Cubic.fm developers found an innovative solution to overcome this obstacle and developed an app for the app studios of major streaming services. This enabled them to stream *licensed* music via music-streaming platforms that make global licensing agreements with record companies and independent artists. The launch of this product in the app studios of streaming services helped them to reach out to more users and better observe their behaviour. Seeing new business opportunities, they consequently launched their stand-alone social media platform: *Cubic.fm*.

5.2.3. Vision and Business Model

Although the company's product-platform changed significantly over time, the fundamental

problem Cubic.fm responds to has remained more or less the same throughout its evolution. At the time of its inception, online listening activity in the music domain was considered to be highly scattered; people were using different services to listen to, discover, and share music. In the Cubic.fm team's view, such fragmentation leads to a lack of data centralization, which prevents music services from offering listeners an effective music discovery experience. Moreover, it pushes artists and record labels to invest great effort into reaching out to potential fans across different platforms. In this regard, the lack of data centralisation was identified as a crucial problem both for users and for the industry, one to be effectively resolved. Figures 5.4 and 5.5 below are extracted from the company's media kit, where the problem domain and the business model are explained.

Figure 5.4 The problem domain explained in the company's media kit.

Figure 5.5. Explanation of Cubic.fm's business model in the company's media kit.

In order to address the problem of fragmentation in music discovery, Cubic.fm centralises music consumption data and builds up comprehensive 'Music IDs' (music taste profiles) of listeners. The core objective of the team is to enable listeners to be able to reach out to other listeners using different services, while simultaneously enabling artists and labels to reach out to or target listeners and fans personally. As a significant part of its data-driven business model, Cubic.fm has no plans to charge end-users. It is rather the *data* users generate that is of value to Cubic.fm. Connecting listeners of different music platforms in one space, the company aims to consolidate data aggregation on music consumption. In this way, it can support and optimise the targeted promotions of artists' and music labels. Offering subscription-based fan engagement opportunities for artists constitutes one of the many revenue streams considered by the team in addition to alternative revenue points, including brand sponsorships and affiliation programs with streaming platforms. In the video presented in Figure 5.6, one of the co-founders explains their business models to a group of investors in a start-up competition.

Figure 5.6 A screenshot of a video about Cubic.fm available on Youtube in which one of the co-founders pitch their product to the investors and afterwards answering investor's questions about their business plan and long term goals. Retrieved from Youtube 'Cubic.fm, Startup Turkey 2015'. Available at: https://www.youtube.com/watch?v=LEvJQGw27wM

5.3. Platform Design and Data Operations

I previously introduced the context in which Cubic.fm was created. In this section, I depict the platform's evolution, describing the design of the three different forms that the Cubic.fm platform has taken and the rationale behind changes made in the development process.

5.3.1. First Version of Cubic.Fm: Crowd-Sourced Radio Stations

Cubic.fm was first launched as a stand-alone website in the form of a crowd-sourced radio service. It grouped radio stations under specific mood/activity categories to thereby offer music suited to listeners' *moods* and *activities*. These stations (playlists) were oriented around particular genres that were created by developers. But the service was described as a 'crowd-sourced radio' because of features that allowed users to make contributions to its radio stations. The term '*cubing'* refers to the functionality through which a user adds songs to the

editorially created radio stations. In this way, the radio stations become crowd-sourced.

This was the initial structure upon which user participation in the first version of Cubic.fm was built. It was later replaced with a different configuration of user-participation, in which users could share songs from their profiles, follow other users, and listen to the songs shared by the users in their personal feeds. The final form of this crowd-sourced radio service was thus built on a combination of the activity/mood-genre-based radio stations – produced through editorial and algorithmic curation - and a specific form of user participation based on a personalized feed structure. I will now describe its features and functionalities in greater detail.

Figure 5.7. Press coverage about the crowd-source version of Cubic.fm. Retrieved from 'All My Favs'. Available at: https://blog.allmyfaves.com/music/cubic-fm-crowd-sourced-radio/

- **Basic Structure**

The Log-in structure of this crowd-sourced version builds on an authentication system based on users' social media accounts. A user could sign-up to the system with her Facebook or Twitter account. After signing up, the system asks for consent to aggregate information from her social media accounts. The structure of the platform is based on three core parts: the *Suggested* page, the *Browse* page, and the *Player* section. The Suggested page is the first page users see after login. Here, users are offered a list of activities and related music genres as subsets of these activity categories. This list is presented to the user in a *personalized* way based on the predictions that the system made on users' local conditions and past listening behaviour.

The logic of personalization in the crowd-sourced radio version was mainly based on predictions that were made by gathering data on the user's local context – for instance: her location, the weather conditions of that specific location, the time of day and the day of the week. This data was used to suggest activities that the user might be doing (or up for doing) at that specific moment. The user was presented with a set of *predefined* activities matching her local conditions from which to make a selection. For example, depending on the local conditions (e.g., a rainy Monday morning or a Saturday evening in the summer), predefined activities could be anything from "working," running" or "partying" to "chilling over cocktails," "dinner with friends," "on a road trip," "on a romantic dinner." Each activity came along with a group of genre categories tailored for the activity (See Figure 7.1 for a screenshot of this page.). It was the Cubic.fm team who pre-determined these activity/mood categories together with a matching set of genres. This is explained further in the passage below:

> *In these radios we use three filtering steps. First, we offer the activities to the user based on the user's local conditions, like suggesting 'you might be doing these activities right now'. Then in the next step, we show them the music genres that might be compatible with this activity. Or rather we show them the radio stations with those genres. When you choose 'I'm running', you get six radio stations that have their own names and specific music genres. When the user chooses that radio, we already determine which songs should be there as a result of our editorial and algorithmic work in the background* (Ozgur 19/04/2014).

Figure 5.8: Screenshot of the activity-genre selection on the platform's Suggested page

For a newly joined user, local context data (e.g. location, time, and weather) were the main source of data used in personalization. But as she continued to use the system, her preferences were also integrated into the personalization process in the form of behavioural data. While showing the user a particular sub-set of genre categories matching the selected activity type, the algorithm was also taking into account the user's past listening data. In this way, the personalization process was built on a logic that brought together the datafied versions of a user's local conditions and her online behavioural past. In the following vignette, one of the developers explains the filtering mechanism behind this:

> *First of all, we come up with 6 activities for users based on their local time zone, weather and the day of the week. I mean, we pick the activities based on the user's local conditions through a pre-matching of the activities and the music genres that would go with them. For example, we offer the users more energetic types of music for running. So we offer an activity to the user, then based on the compatibility with this activity, we come up with different music genres. Once we've chosen that activity and the genres, then we determine the content deciding which songs would go with them.*

> *But we also consider the choices that the user makes over time. Personalization comes into the picture at this point* (Ozgur 19/04/2014).

User participation in this crowd-sourced radio version was built on the 'personal feed' structure. In this structure, users could *follow* other users to see their activity in their personal feed. This may include sharing a song ('cubing') or listening to particular radio stations. In order to 'cube' a song, a user needed to search and select a song from the Cubic.fm database, enter her personal message and tags, and then click on the "cube" button. Then the cubed song was queued in the personal feeds of the users who were following this user. The Newsfeed of the micro-blogging platform, Twitter, inspired this structure. It was developed to facilitate music discovery through the songs shared by other users. Listeners could interact with the 'feed items' by "cubing" (sharing) a song; "re-cubing" (re-sharing) a song shared by someone else; "liking" a song; or commenting on a shared song (See Figure 5.9). In this regard, the currency of interaction in this version was single songs while the means of interaction was the personal feed.

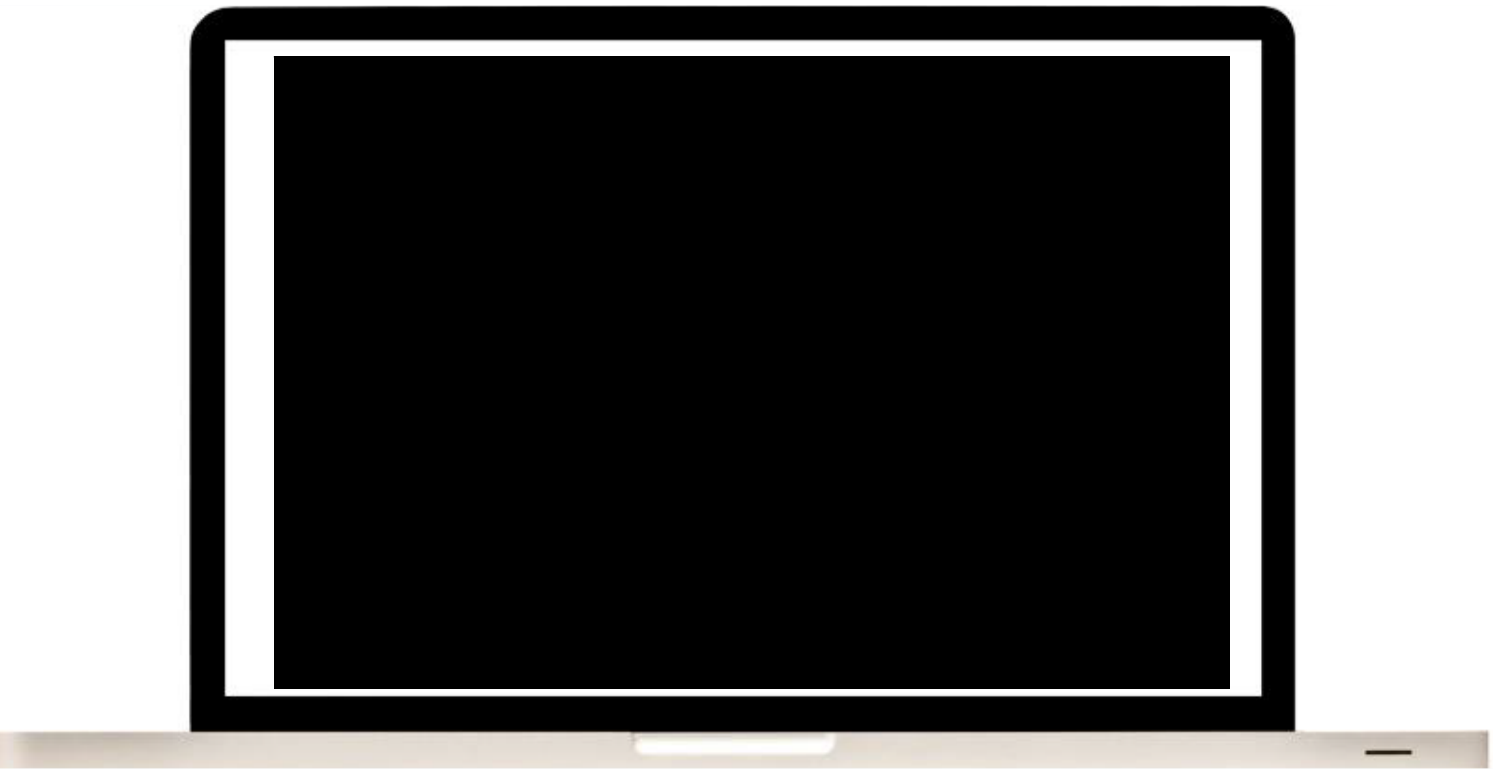


Figure 5.9: Screenshot showing 1) the Player on the top left (the activity & genre station on the player (now playing) is "working" & "downtempo") and 2) the personal feed structure at the center of the page.

One of the key features of the Cubic.fm platform in its first iteration was the public presentation of listeners' 'Music IDs', in which each user's interaction with music was

recorded by the platform and visually presented to other users. It is worth noting that the visual version of the music ID here is *automatically* generated by the system based on continual behavioral tracking, as opposed to the bio or profile page created by the user. These 'Visualized Music IDs' include behavioral metrics, like the number of times the user "cubed a song", "liked", or listened to music on the platform. For example, different colors were used to signify the musical genres engaged with most by a specific user (See Figure 5.10 below). These IDs were subject to change as users continued to interact with music. To illustrate, a user could be presented as having a 'mastery level' on a particular genre based on her past song-sharing (cubing) activity, which equates the contribution of content to the platform to some form of expertise on the genre. Both the metrics and the wording used here are gamification elements that the developers utilized to engage users in participation. I will explore the use of gamification on Cubic.fm in more detail in subsequent chapters.

Figure 5.10: Screenshot of an example of a users' 'Music ID' as a visualised representation of the genre preferences and behavioral metrics of the user.

- **Backend Operations**

These front-end developments were made possible by a set of back-end operations. At their foundation, Cubic.fm radio stations were playlists that were created based on a *taxonomy*

developed by the Cubic.fm team. These playlists were built on an initial set of songs. This initial set was composed of particular songs that were specifically handpicked by the team in order to ensure that their genres correctly match up with the activity-genre category of the related playlist (radio station). Although initially created as static playlists, over time these playlists were turning into dynamic radio stations. The system was continuously aggregating music and its metadata on the web to automatically add new songs to the existing stations. Therefore, the initial set of editorially selected songs was serving as a reference point for the categorization of the newly added songs.

> *First, we create the identities of these radios with our own work without getting much algorithmic support. It's because we have very little data when we first start. So let's say there are to be 50 songs to be played in a radio. We first determine these 50 songs, and then begin to develop the radio's identity algorithmically. But we are the ones determining the core of them all by using our own musical experience or by following music blogs and alike on the Internet. Only after we form this core part ourselves, we start to algorithmically expand these radios* (Ozgur 19/04/2014).

i. *Editorial Work behind the Radio Stations*

The editorial work underpinning the radio stations was based on a taxonomy developed by the Cubic.fm team. It was used to link predefined activity/mood categories to the supposedly related genres. This taxonomy was developed based on the experience of the team in developing music recommender systems, which requires extensive knowledge on ordering music and transitions between different genres. This process was explained by one of the developers in the following vignette:

> *When we define an activity on Cubic.fm, we determine in which mood and genres could there be playlists in this activity. When we determine (define) an activity, we rely on a taxonomy that we developed ourselves. So that when an activity is defined on Cubic.fm it would mean that it has got certain tags. When a user selects an activity, we connect it with the user's Music ID through these tags* (Onur 08/05/2014).

In order to ensure the reliability of the taxonomic structure they developed, the team was using extra sources, such as services like *The Echo Nest, Last.fm* and *Beatport* etc. as explained below*:*

For example, now we go to the Echo Nest, and it could give us the relevance of a tag with another one. That's why we can see how the tags exist as clusters. Also, we can see the similarities between genres through the data we crawl on the Internet. For example, Indie-Rock has been most frequently tagged with the tag "rock", and we can see that. Or "minimal" has been tagged most with "techno", and we can see that as well (Onur 08/05/2014).

ii. ***Aggregating Music Content***

Cubic.fm was continuously adding new songs to its database through aggregation. In this way, content production could be done algorithmically on top of editorially created stations. In the following vignette, a developer explains how a song enters into the system:

So how does a track enter into Cubic.fm's database? That's the most important part. Now we're crawling a lot of resources; these could be the playlists on Spotify, Deezer, Songza or Pandora. Name a service that provides playlists or radio stream, we are most probably crawling it. It means we collect its track data. We collect them with their own APIs or with the scripts that we've written. We primarily collect playlists. Collecting playlists rather than single tracks is much easier. It's much easier to identify and categorize music this way. Once we pull in a playlist, we start identifying and categorizing each song in it. Some of the tags of the tracks come from the tags of the playlist. I mean when we retrieve a playlist from Spotify or Last.fm, it has tags right, we can assume that these tags are also the tags of the tracks in it. The logic is the following: if I call a playlist 'happy', then there's a certain possibility that the songs in it are also in a happy mood (Onur 08/05/2014).

iii. ***Aggregating Metadata***

As explained above, after aggregating music content, the system needed to identify, analyze and categorize the music tracks. In order to do so, the system was also continuously aggregating metadata of the content from different sources. There were two main sources for that: *user-generated tags* and what the team calls *the scientific metadata* about music tracks.

The Echo Nest and Last.fm were the main sources for aggregating metadata for ordering music. Keywords or tags generated by Last.fm users to categorize songs were continuously aggregated and kept up-to-date using Last.fm's API. Those user-generated tags were

analyzed in combination with the scientific data acquired from the Echo Nest, which provided mathematical values to describe certain qualities of music objects. These data points were aggregated using the Echo Nest's API (See Figure 5.11 for a developer's sketch on these data points). One of the developers explains this as follows:

> *We also have the data that gives us information about the song. Let's say it's the science of music. The Echo Nest provides us this scientific data. We were actually creating this data ourselves using signal processing techniques but then we realized that the Echo Nest had already prepared for all songs the values that are very much similar to the values that we determined. Therefore, we didn't have to do such things like signal processing for each and every song because there was already available data. From the Echo Nest, I get the song's bpm (beat per minute), tempo, danceability, loudness, how it changes from which decibel to which one, etc.… So the Echo Nest shares with us the data it gets from the song's signal, it's actually pretty scientific. These are the values that their (The Echo Nest's) research team determines… We have these scientific values. They are all numbers; bpm, for example, would be something oscillating between 60-150. Danceability changes between -1 and 1. Loudness is always a decibel (db). They are all mathematical values* (Onur 08/05/2014).

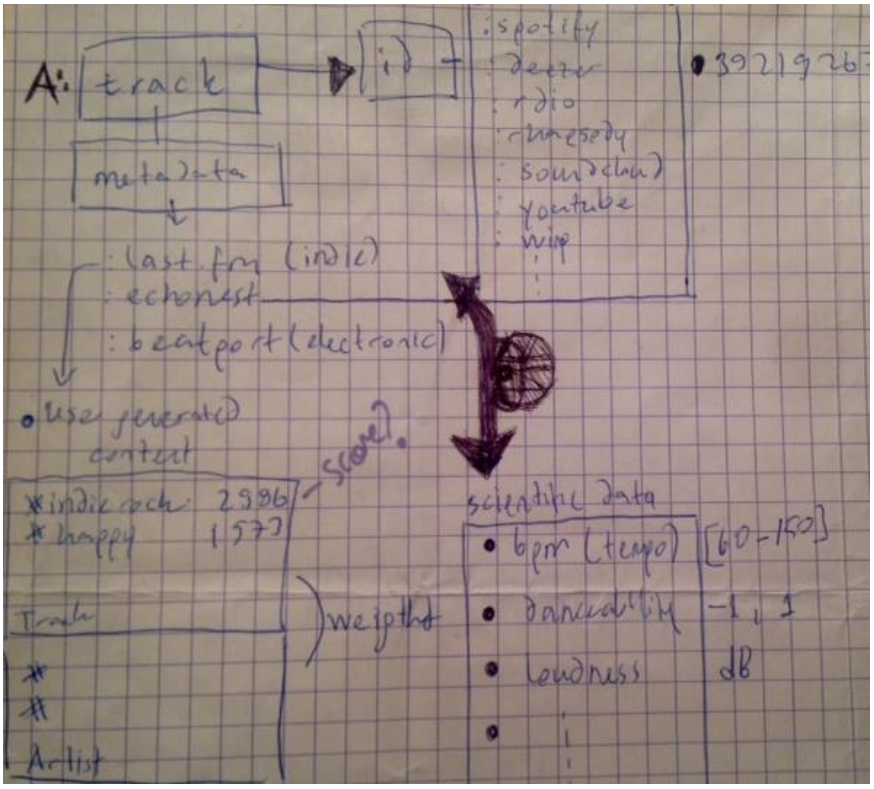


Figure 5.11: Sketch of a developer showing values for the scientific data points kept for the tracks (see the extract from Interview with Onur 08/05/2014).

iv. *Ordering Music Algorithmically*

Ordering music items is one of the core back-end operations upon which the entire system of Cubic.fm is built. In the first version, Cubic.fm developers were using machine learning (clustering) algorithms, like neural networks, in order to automatically categorize music tracks using the data aggregated from the Echo Nest. In the process of training these machine-learning algorithms, developers were also using user-generated tags aggregated from Last.fm. These tags were used to supervise the algorithm. But, the user-generated tag data, describing a particular aspect of music like its genre or mood, was only used after being verified as reliable categorical data. In the following vignette, a developer explains this process (See also Figure 5.12 for the developer's sketch explaining the process).

> *So what does this give us? The machine learning comes into play here. Once we're certain that a song is "indie-rock", we mark it as "base". We say that this song deserves to be stored as indie-rock in our database. Let me put it in simpler terms: Let's say we've accepted 100 tracks as indie-rock. For this cluster, we can get a lot of stuff (statistical value) such as mean, median etc., from the scientific data that comes from the EchoNest. Then a track comes enters into my system, and, the user-generated content (aggregated from Last.fm) tells me that the genre of this song is indie-rock. I can understand whether this song is actually indie-rock or not. I do this by checking if this track fits into this cluster or not. So, let's say this cluster is called indie-rock, and the track's name is B. I calculate a distance between track B and the cluster "indie-rock". So if this distance is less than or equal to the threshold, then I can begin to accept this song as indie-rock* (Onur 08/05/2014).

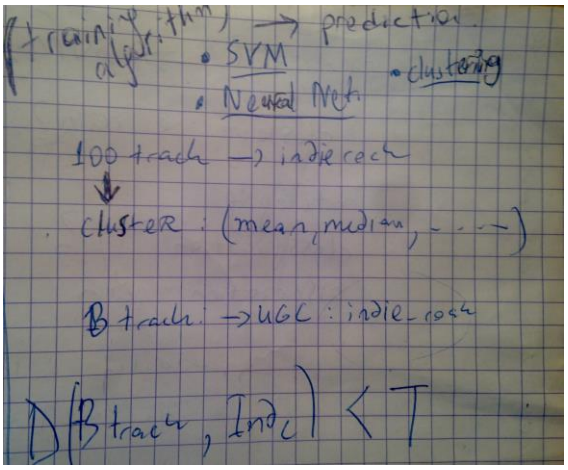


Figure 5.12: Sketch of the developer explaining this process. T stands for the Threshold.

5.3.2. Second Version Cubic.Fm: Social Media Platform for Music Discovery

The structure of the second version of the Cubic.fm platform was notably different from the first version. In the new version, the service turned into a *social media platform* that completely relies on user-generated content. This time it was launched as "a social music discovery platform" aiming to eliminate the boundaries between the music streaming services. The main difference between Cubic.fm as a social media platform from the original crowd-sourced radio stations version was a shift from the 'editorial and algorithmic' curation of content to 'user-generated' content curation. This also changed the currency of user interaction from individual *songs* to user *playlists*. The new version came out as a playlist-sharing platform whereby users of different music services could discover new music through the playlists created by other users with similar music tastes. The logic behind this version is explained the company's newsroom in Figure 5.13 and in a presentation for a start-up competition in the video provided in figure 5.14 below.

Figure 5.13: A Newsroom article explaining the new version. ‘Cubic.fm lets you share and discover playlists, via any streaming service’. Available at: https://cubic-fm.pr.co/75240-cubic-fm-lets-you-share-and-discover-playlists-via-any-streaming-service

Figure 5.14 ‘Music Discovery is Still A Problem! But Why?.Here is Cubic.fm’. Available at: http://blog.startupistanbul.com/tag/cubic-fm-startup-istanbul-presentation/. The Video content can also be reached at Vimeo: https://vimeo.com/108549862

- **Basic Structure**

The basic structure of the system was based on eight basic modules: Home, Explore, Search, Profile, Notifications, Settings, and Library and the Player. The main functionalities were built around playlists (creating new playlists or importing, editing and sharing existing playlists). The Login structure of the platform was again built on an authentication system based on users' social media accounts, so that the system could aggregate data about users upon signing up the system. Thus, it starts building user profiles from registration to eliminate the 'cold starter' problem.

Users could create new playlists either on Cubic.fm platform or by *importing* their *existing playlists* from on other music services. They could create a playlist on Cubic.fm or edit existing ones by adding songs to them from other users' playlists available on the platform. The 'Add' button appearing next to the songs playlists enabled users to add this song to their own playlist (See Figure 5.15). Alternatively, they could import and edit existing playlists by connecting their streaming service accounts to Cubic.fm. Users were given the option to save playlists as drafts or share them to with other users publicly. In order for a playlist to become sharable it needed to fulfil certain criteria. This process was called the "*Road to a Perfect Playlist.*" In order to become eligible for publication, a playlist needed to have: (1) at least ten songs in it, (2) a title naming it, (3) a cover photo visualizing it, and (4) at least three tags describing it (See Figure 5.16 and 5.17). Playlists failing to match these criteria remained as draft playlists that were invisible to other users.

Figure 5.15: An example screenshot showing the “Add” button

Figure 5.16: Screenshot of the playlist creation process

Figure 5.17: Example screenshot showing the process of adding song to a playlist

The Home page featured *personalised* suggestions composed of other users' playlists, as well as suggestions for trending, popular, and newly added playlists (See Figure 5.18). The Explore page provided users the opportunity to browse user-playlists by selecting 'tags' (See Figure 5.19).

Figure 5.18: Example screenshot showing how personal recommendations are presented.

Figure 5.19: Example [redacted] showing the Explore Page

In the Search page, users could look for particular *tags*, *users* or *playlists* available on Cubic.fm. The "Your Music," section featured the collection of users' own playlists (imported, drafted and/or published), the playlists they were following, or the playlists they had recently listened to (See Figure 5.20). The "Player" allowed a user to listen to a playlist. She could play, skip, and pause a particular song, while she could also 'like' it, or add it to one of her playlists.

Figure 5.20: Screenshot of the "Your Music" Page.

Personalized suggestions made on the Home page were based on the similarity between the ‘music taste of the user’ and the ‘playlists’ generated by other users. User-generated playlists were considered to be the best indicators of one’s music taste. Therefore they were taken as the basis to facilitate music discovery among users. Based on the behavioural history of listeners (i.e. what they listened to, liked, or followed before) users were recommended playlists created by other users that involved songs with their favourite artists or the tags the listeners interacted with most (i.e. ‘liked’ or listened to songs with the same tags, etc.).

The means of *interaction among users* was the ‘Follow’ functionality. Users could follow a particular user; go to her profile; see her playlists (recently played, followed, and published); see the number of users following her and the number of users followed by her. Users could also follow a particular playlist to receive notifications about it (i.e. when a new song was added to it). Additionally, users could *interact with playlists* by ‘liking’ songs in a particular playlist, commenting on a song, adding a song from that playlist to one of their own playlists, or sharing playlists on social media platforms like Facebook and Twitter. Each playlist had a ‘feed’ section showing the latest activity on the playlist (i.e. users who recently started listening to or following the playlist, ‘liked’ a track on the playlist, etc.). Right above the feed structure, users were also provided metrics regarding the activity on the playlist (i.e. the number of times the playlist was listened to, the number of its followers, and the number of activities on the playlist (See Figure 5.21).

Figure 5.21: Example screenshot showing playlist metrics on the right of the playlist cover photo.

Furthermore, the system provided users with notifications about the playlists and the users they followed. This could be when a new song was added to the playlists they were following or when a user they were following created a new playlist. For example, users also received notifications when other users interacted with their published playlists. They were notified when a user started to follow one of their playlists, liked a song in one of their playlists, or added a song from these playlists to their own ones (see Figure 5.22). Lastly, users could receive notifications via email (Figure 5.23 and 5.24) and/or manage their settings (See Figure 5.25).

Figure 5.22: Example screenshot showing the notifications

Figure 5.23: Example screenshot showing email notifications

Figure 5.24: Example screenshot showing an email notification

Figure 5.25: Example screenshot showing notification settings

- **Backend Operations**

The structure of user participation on the second version of Cubic.fm platform described above was supported by a set of back-end operations. Given that the core value proposed to users in the new system was the facilitation of music discovery by matching users with relevant playlists suited to their tastes, back-end data operations were of crucial importance in the new system. Cubic.fm did not use demographic user data to personalize content, but rather took users' 'interaction with music' as the basis from which to give personalized recommendations. Therefore, the logical structure of the system was built on the *decomposition* of 'interaction with music' to produce taste profiles of listeners and thereby match them with relevant playlists. Matchmaking between listeners and playlists was done by aggregating and centralizing various data points from different sources (See Figure 5.26 and Figure 5.27 for the lists of data points according to source platforms). These data points were used to build users' music taste profiles referred to as 'Music IDs'. In the following section, I explain this process with the logical structure of the database.

i. Building the Music ID

In order to solve the 'cold-start' problem of first-time use, Cubic.fm built its login structure with a sign-up process based on social media accounts. When users signed-up to Cubic.fm using their social media accounts, the system imported users' previous music activity data (e.g. shared videos, 'liked' songs and artists etc.) using the social media platform's API. In addition, users were also given the option to connect their streaming service accounts to Cubic.fm. When they connected their streaming services, Cubic.fm pulled in user's playlists together with previous behavioural data such as 'liked' songs and albums, followed artists, listened to songs, etc. Thus the initial Music ID of a user was built. After this initial stage of data aggregation, the Music ID of the user continuously developed as she interacted with music on Cubic.fm platform.

Figure 5.26: Screenshot of the document showing the data points and their sources. This document, provided by the company, shows the data points according to their source platform in the form of a matrix composed of action-object relations - objects are the entities (i.e. users, songs, playlist, artists etc.), whereas the actions are the activities users can perform on the platform. On the left side of the Excel sheet is the generic list of possible action-object relations that can be found in other platforms. The right side of the sheet shows which action-object relation could be available to aggregate on which platform (Here SS stands for streaming services, FB for Facebook, TWTR for Twitter, SC for SoundCloud and LAST for the Last.fm.)

Figure 5.27: Screenshot of a different representation of the same data points and sources in Fig 5.26, but in this figure, data points are listed according to the source platforms separately.

The process of decomposing users' interaction with music includes a set of steps. First, this high-level concept called 'interaction' is broken into discrete 'action-object relations'. To illustrate, although playlists are the primary currency of interaction, all sorts of user activity data aggregated from social media, streaming services, and on Cubic.fm's own platform are being *normalized* in Cubic.fm's database as discrete 'actions' permitted on the platform in relation to the core 'objects' of Cubic.fm system: *Tracks*, *Artists*, *Playlists*, and *Tags*. The resultant action-object relations could be anything a user can do on the platform, from 'liking-song' to 'creating-playlist'. This is basically a process of *normalization*. Different types of actions specific to different platforms are being normalized during the data aggregation process. Thus, different types of actions signifying similar activities are being standardized and stored in a unified form in Cubic.fm's database. One of the developers explains this in the following vignette:

> *The data points that we get from other services change from service to service because every service has its own data points. For example, you cannot like a song on Rdio so we cannot take 'like' data from Rdio… Or we cannot get playlist data from Last.fm because it doesn't allow you to make playlists, but we can get playlists from Spotify because it does allow it. On Deezer, you can have a favourite album, but I don't think you can on Spotify. You can follow playlists instead. What we are doing is actually… whatever you can do with an album; let's say listening to it, making it a favourite one, following it, or liking it, when we get these data points inside, we*

> *normalize them. Similarly, you can add an artist to your favourites on Spotify, like an artist on Deezer, and follow an artist on Rdio. What we have at hand is three different types of very significant data. But their names are different. They're all going to mean the same thing for us. And this is essentially unification* (Onur 18/07/2014).

As a second step, action-object relations are further decomposed in the form of a user's relation with the 'tags-of-the-objects' included in the 'action-object relations'. In this way, the system is built on a particular logic that sees each action of a user in connection to the '*tags'* of the '*objects'* of these 'action-object relations'. For example, let's say a user, Mary, 'likes' a song from Bob Marley. The song is called "No Woman No Cry". The tags of this song are "Reggae," "Jamaica," "Rastafarian," "70s," etc. As the first step, Mary's interaction with music is decomposed and standardized by turning it into an action-object relation. The action-object relation in this interaction is the following: 'Mary-likes-NoWomanNoCry'. In the second step, extracting the tags of this song and relating it to the user in question normalizes this action-object relation. The resultant normalized version becomes the following user-tag relations: "Mary-Reggae"; "Mary-Jamaica"; "Mary-Rastafarian"; "Mary-70s." These user-tag relations are stored in the database as data objects that can be *scored* and *ranked*. As a result, a number of 'user-tag scores' are calculated for each user and these scores are taken as datafied representations of the users' music tastes. This process is repeated for *each* tag of the objects involved in the action-object relations that are created in the course of users' interactions with music on the platform.

In the calculation of 'user-tag scores', the system weighs the *type* and the *recency* of actions *differently*. Explicit actions are considered to be more accurate indications of a user's taste. These are also the kind of user actions desired and encouraged on a social media platform. So, the more recent and explicit the actions are, the higher they are weighted. For example, pressing the 'like' button for a song in a playlist was considered to take more effort than listening to a playlist. In other words, listening is perceived to be a more passive action than explicitly 'liking' a song in it. Therefore, the action of 'liking' a song is weighted higher than listening to it. In the same vein, tagging a playlist weighs more than liking a song in it, because tagging requires pressing the tag button, typing some keywords, and saving them. Consequently, the higher the level of the effort an action requires from the users, the higher the action is weighted.

As a result of this process, calculated tag-scores are used to rank the tags that a user engages

with most. Thus the top ranking tags guide the system to find the playlists matching the user's taste. In order to find matching playlists, Cubic.fm calculates a similar tag-score for each playlist available (created or imported) in its system. The calculation of tag-scores involves aggregating the metadata of the tracks and the artists detected in each action-object relationship. Last.fm and the Echo Nest are the mostly used services as sources of metadata for this. Playlists that match a user's taste are identified based on the calculated similarity between the highest scored tags of the user and the playlists. After identifying relevant playlists, a personalized list of suggested playlists is produced and presented to the user on her home page. Users' *reactions* to these suggested playlists (e.g. starting to listen to a playlist etc.) become the *new data points* to be normalized and scored. Thus, users' music tastes are *continuously updated*.

ii. Logical Design of the Database

The logical design of Cubic.fm's database is built on action-object relations and their normalizations. This is accomplished by creating relational entities. The core entities (objects) of the system are: *User, Playlist, Track, Artist* and *Tag*. These are thus the main tables in the database. The rest of the tables that we find in the Entity Relation Diagram (ERD) of the system are the relational tables that enable normalization (See Figure 5.28 for the original sketch of the developer.).

The key table in the database is the (User) Tag-Score table. This is a relational table connecting 'User' and 'Tag' entities. Each Tag-Score object created in that table is composed of a 'userID', a 'tagID' and a 'Score' calculated for a specific user for a specific tag. Every time a user takes an action, for each and every tag involved in the action-object relation created for this activity, a new object is created and inserted in this table. If there is already an existing object for a tag, then the score of that tag is updated. Thus, a quantitative representation of the user's music taste is produced and continuously updated. Although a user's taste manifests in the form of "the highest-ranking tags" in the Tag-Score table, the production of that score heavily relies on the relational tables used in the process of normalization. Relational tables like *Track-Like, Artist-Like, Playlist-Follow, Playlist-Track, User-Follow* are all created for the action-object relations. Likewise, relational tables like *Artist-Tag, Track-Tag* or *Playlist-Tag* are used to identify the tags of the object of these action-object relations. In this regard, quantifying users' music tastes is accomplished through the normalization of user actions through these relational tables.

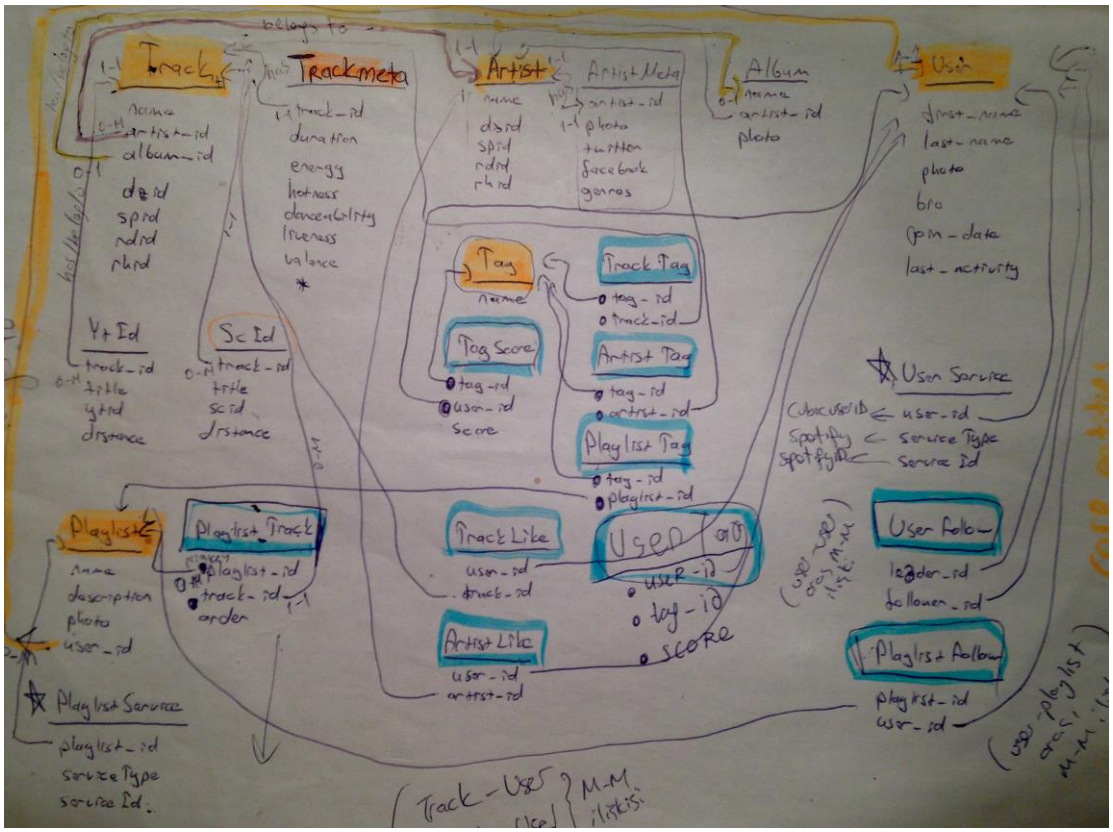


Figure 5.28: Image of the original sketch of the ERD by the developer

5.3.3. Third Version of Cubic.Fm: Online Music Library

- **Basic Structure**

The latest (and the current) version of Cubic.fm was launched as an *online music library* connecting different music sources in a single player (see the demo video presented in figure 5.29 below). In the new system, users are provided with a browser extension that enables them to save audio content while they are browsing the web. When a user downloads the Cubic.fm extension, the system embeds a "Save-to-Cubic.fm" button under the audio/visual posts on the web sites that the user visits. This 'save' button is embedded in those websites in line with their native design. In other words, when, for instance, a YouTube video is shared on Facebook, a "Save-to-Cubic.fm" button appears underneath the video on Facebook right next to Facebook's 'like' button (See Figure 5.30 and 5.32 for the example screenshots).

Figure 5.29 Demo Video available on YouTube explaining the Personal Library Version of Cubic.fm. Retrieved from Youtube 'cubic.fm Demo Video'. Available at: https://www.youtube.com/watch?v=qd7cGUSKMRM

Figure 5.30: Screenshot of the "save-to-cubic.fm" button that appears on SoundCloud after downloading the Browser Extension

Figure 5.31: Screenshot of the “save-to-cubic.fm” button that appears on Facebook

When the user wants to save a song or a video she comes across on Facebook, she clicks on the “save-to-Cubic.fm” button. Then a pop-up window appears on the right-hand side of the screen asking the user to select the playlist that she wants to add the song to (See Figure 5.33). Alternatively, the user can also save the ‘URL’ of an audio post by copying and pasting it manually in the ‘add music’ window on the Cubic.fm platform (See Figure 5.34). Users can also categorize the music they collect across the web by putting them in different ‘channels’. These channels are basically dynamic playlists that users can create and expand as they discover new songs. These playlists can be named and categorized at the same time by simply adding tags while they are created and/or edited (See Figures 5.35 and 5.36).

Figure 5.32: Screenshot of the pop-up window

Figure 5.33 Example screenshot showing how to add songs using the ‘add song’ feature on the platform.

Figure 5.34: Example screenshot showing how users can categorize their channels (playlists)

Figure 5.35: Example screenshot showing the channels (playlists) and the tags describing them

- **Backend Operations**

The browser extension launched in the latest version of Cubic.fm is a "front-end heavy" tool, as described by one of the developers. When downloaded by the user, the extension asks the read/write permission from the user in order to embed the "save-to-cubic.fm" button on the web pages that the user visits (See Figures 5.37 and 5.38). The prominent back-end operations is composed of: 1) a process of aggregating music items from different sources; 2) then identifying and matching the same songs aggregated from different services; 3) and finally assigning them a unified ID -similar to ISBN of a book. The database structure is based on standardizing and unifying the music ID's of the aggregated songs in order to find them on other services and stream them in a source-independent way. A team member explains this as follows:

> *We started to look at it more simply. So here is the idea, every song is actually a link inside a service. We should put a code to that song, a barcode-like number. Once we*

are able to put the same code to the same song that might appear in other services with a different name, then we can see where else we can find the same song when we want to stream it. This way when a user saves a song from Spotify we can stream it from free services like YouTube or SoundCloud **(**Baris 18/12/2015**).**

Figure 5.36: Example screenshot of the ChromeWebStore showing the read/write permission asked while downloading the browser extension

Figure 5.37: Example screenshot showing the explanation available on Chrome Web Store regarding how Cubic.fm browser extension works

The process of building a user ID is also simplified and made more flexible for new functionalities that will/can be added in the future. In this regard, one of the developers compares the construction of user tag-score in the previous version and the new system as follows:

In the previous version, the product was a little more complex. In the process of calculating tag-scores, first there was the data coming from Spotify, Deezer etc. This was the behavioral history we were able to extract from these services. Second, the

user could like a song, add songs to her playlist, listen to songs on Cubic.fm etc. There was data coming in from different sources. And calculation of tag-score was relying on a complex algorithm. Contribution (weight) of adding a song to a playlist to the user tag-score was different from 'liking' a song etc. Now in the latest version, you can imagine it in a much simpler way. We now only aggregate data in two ways: 1) when a user 'saves' a song and 2) when a user 'listens' to a song on Cubic.fm. Using just these two data points we can get a much simpler version of the exact same tag-score structure. But of course as we put more features to the product, we would be able to collect different types of data in the future **(**Ozgur 18/12/2015).

As the vignette above illustrates, in its latest version the Cubic.fm platform has taken the form of a *personal* music library without any social features. The main functionalities of the system now are the browser extension and the dynamic user playlists. In this last version both the frontend features and backend operations are significantly simplified. However, the team still has a roadmap to further develop the platform by adding new functionalities in the future, hinting that the Cubic.fm will continue to evolve. The next section describes the process through which the platform evolved up to this point, from a crowd-sourced radio to a social media platform and then finally to a personal library.

5.4. Development Process

In this section, I will outline the *evolution* of Cubic.fm's product-platform. In doing so, I will describe the main reasons behind the changes made in the platform functionalities which effectively brought about its redefinition as three distinct versions over time.

5.4.1. Evolution of the Crowd-Sourced Radio Version

Cubic.fm was first launched as a static radio app based on passive music listening. It then evolved into a stand-alone crowd-sourced radio platform with social features that allow users interact with each other. The form of user participation in this version was built on two core features, namely 'cubing' (sharing) a song and the "feed" structure. These social features were tested in a private beta for three months. In this period, several problems were observed with regard to the structure of user-platform participation. To address these problems, the team decided to abandon the structure built on individual song sharing and instead move towards a new one in which the currency of social interaction is user playlists.

One of the most prominent challenges that led to this change was associated with finding reliable metadata for music. As discussed in detail, the crowd-sourced radio version was built on editorial and algorithmic classification of music. Hence, the service was heavily reliant on music metadata provided by third party platforms like the Echo Nest and Last.fm, which the team called the 'metadata services'. Since metadata play crucial roles in the algorithmic analysis and ordering of music, developers needed to make sure the data were reliable. One of the developers explains this challenge as the following:

> *We tried a number of different methods, having the algorithmic and the editorial together, to see what works best in order to offer better radio stations. I'll give you a very simple example: I don't remember which radio it was but it actually didn't have that good of an accuracy rate because of a wrong method we used for detecting the genre of the song, or rather we over-trusted the data that the service provided us with* **(**Ozgur 19/04/2014).

To address this problem, the team developed a solution based on building a trust relationship with the metadata services, especially in the process of aggregating user-generated tag data about the songs. The following vignette explains this process in detail:

> *Last.fm gives us the number of times a particular tag is used for a particular song. Let's say we have a track named A, we know that this track A has been tagged as "indie-rock" 2986 times, and tagged as "happy" 1573 times. We take this tag data. The track also has an artist. Similarly we take the tags of this said artist, too. Then we put a weighed value to it within our algorithm. I mean, just because a song has been tagged "indie-rock" a certain number of times doesn't mean that it is an indie-rock song. That's why we determine a weight for every value we take from Last.fm. What the weight refers to is the number of times it has been tagged. Here we also look at how many times a song has been scrobbled and tagged on Last.fm. The conversion in-between has something to tell us. Also, we try to get a relevance by looking at how many times a track's artist has been tagged with exactly which tags. In the end, we identify a threshold for a track this way. The weight here shows to what extent we are to trust this tag.*
>
> *[...] In the end we came to such a point that, for example, in the APIs that we use there's also Beatport, which is a great API for electronic music, but we could not find*

> *anything about indie-rock. Last.fm is terrible for electronic music but pretty good for indie music. So if Last.fm gives the indie tag in the data coming for our track, we're more likely to trust Last.fm. But if it gives the electronic tag, then we're less likely to trust it. Then the electronic tag coming from Beatport becomes more significant for us … but this weight is not a static value, it updates itself. How does it do that? When I play a song where it's tagged indie-rock, the reaction that it gets, whether it's skipped, or liked; and, if it's skipped, was it at the 5th second or at the 120th? Looking at all of this helps us understand if it's really indie-rock or not from the user's perspective. This is how we build a trust relationship with the APIs that we speak with and see how much we really trust them* **(**Onur 08/05/2014).

The process described above exemplifies the work required to aggregate metadata reliably in order to use it in the algorithmic analysis and ordering of the songs. This was one of the core challenges of the crowd-sourced radio version and constituted one of the reasons behind the shift towards the user playlists as more *reliable* and *scalable* way of ordering music.

The design of the crowd-sourced radio version was significantly inspired by the Newsfeed structure in social media sites like Twitter and Facebook. In other words, the Cubic.fm team first tried to build, to use their words, the "Twitter of Music." This version was tested in a private beta for three months. During this beta process the team observed that the Twitter-like feed streams, enabling users to interact with the songs shared by other users, was not effective for creating efficient levels of user engagement. Most of the problems spotted in the feed structure were related to the particular nature of music content and the associated consumption experience specific to this kind of cultural object.

First, it was observed that presenting users with shared songs in a time-based sequence in the form of a feed stream creates an exhaustive list. This overwhelms users and decreases users engagement. Users cannot listen all the songs in their personal feed and fail to discover new songs effectively.

> *The feed structure we built was based on users giving feedback to each other on one single song. But we realized that it is very difficult for users to see the whole content and give feedback to all of them* **(**Ozgur 19/04/2014).

Moreover, it was observed that music consumption is not a 'one-time only' interaction. Unlike reading a tweet, the same song could be listened to several times. Therefore, it should

be easily findable. The feed structure that prioritises the most recently shared songs fails to facilitate such an experience. The following two passages explain this challenge together with the solutions the team sought to address them.

> *We observed that the feed was not flowing as fast as a Twitter feed. We weren't able to have the user consume the content the way we wanted to. That was the biggest problem. This is because music isn't something that depends so much on time. For example, a tweet from a week ago can get lost on your feed but you can listen to the same song even after a year* **(**Erdem 05/06/2014).

> *[…] I cannot expect the user to sit in front of the screen and listen to the songs by scrolling down and clicking on the songs because music is just not consumed that way* **(**Erdem 29/04/2014).

Finally, the sequence of songs shared and streamed on the feed was not following a particular logic, and the songs played subsequent to each other could be in completely different genres and moods. This is observed to create a disruption in the flow of listening. Listeners needed to skip the irrelevant songs manually. Considering music listening as an activity that usually accompanies other activities in the background, the feed structure appeared to be dysfunctional for music discovery. Taken all together, these concerns regarding the single song feed stream posed a challenge for Cubic.fm.

With all the observed problems and challenges related to facilitating music discovery through a feed structure considered, the team decided to reconfigure user participation on the platform. They turned towards user playlists as an efficient solution to the observed problems for the following reasons. First of all, playlists are considered to be better indicators of one's music taste than the single songs they share. Secondly, user-generated playlists are seen as reliable sources for metadata. Thirdly, playlists are more likely to provide smoother song transitions. This is because they are grouped and categorized by humans in a particular sequence and according to a particular mood and genre that can accompany a particular activity.

The idea behind this shift initially was to use the playlist just to facilitate social interaction among users as an added value on top of the editorial and algorithmic content produced by the team. So, in the new version, rather than interacting through individual songs on a personal feed, the idea was to enable users to interact with each other through the playlists

they shared. Thus the feed structure was abandoned and replaced with user playlists. In order to do this, the team decided to mobilise users' existing playlists that were created on other music services. This was achieved by enabling them to import their playlists into Cubic.fm. They initially started to implement this structure by integrating the core streaming services, namely Deezer, Rdio, and Spotify. However, especially after the Spotify integration, the team realized that building the system entirely on user playlists could have several more advantages than just facilitating social interaction.

First of all, user-generated playlists were more meaningful, reliable, and scalable means for content production and categorization compared to the technical work underpinning editorial and algorithmic curation. Secondly, in addition to the aggregation of existing user playlists, streaming services also allowed the aggregation of the behavioural history of the users. The team noticed that this would enable them to centralize behavioural data in one place. Hence it could pave the way for developing a service that eliminates the growing fragmentation in the digital music landscape by connecting users of different services in one place. One of the developers summarizes this process as follows:

> *We knew that creating value through music with data would be valuable sooner or later. And our approach was to try to reach different people by centralizing the very many services that people use. First we started with Deezer. Then when we integrated with Spotify, we realized that both Deezer and Spotify are completely different platforms, but we could unify the users of both of these platforms with one single application. Thus we could have the users of both platforms communicate with each other… We already had our application in Spotify. When you go to this application on Spotify, you can listen to a song added from Deezer by someone else. We saw the connection this way. Then we ran Cubic.fm app in two different places and we actually saw that the data was gathered in one place, a centralized place, that was our platform. Thus we began to better understand how important and valuable this centralized data was. Then it hit us that this was something that people needed. The moment we realized that opportunity, we started to focus on that.* (Ozgur 18/12/2015).

Having observed the benefits of user-generated playlists, the team came up with the idea to completely abandon the editorially and algorithmically curated radio stations and instead just build a social media platform for music discovery around user playlists. The data-driven vision of the company thus further crystalized as to create value both for the users and the

industry through the elimination of fragmentation in the digital music landscape. The goal was to enable users to reach out to the playlists of the users of different streaming services while centralizing the behavioural data aggregated on the streaming services. Consequently, the new Cubic.fm was launched in 2014 as a social music discovery platform.

5.4.2. Evolution of the Social Discovery Platform Version

The second Cubic.fm version was launched in April 2014 as a *platform-agnostic social music discovery service* connecting users of different streaming services via user playlists. Initially partnered with Spotify, Deezer, Rdio, Rapsody, YouTube and SoundCloud, the company's goal was to centralize the music listening experience and music consumption data that was fragmented and dispersed across different music streaming services, and thus to build users' Music IDs.

The idea behind this was to help users reach out to people with similar tastes through their personal playlists no matter which music service they were using. Thus the value proposed to users was to eliminate the boundaries between different services. To illustrate, this would allow a *Spotify* user to be able to access and listen to a *Deezer* user's playlist through Cubic.fm. Standing in the middle of different streaming services, Cubic.fm would facilitate music discovery by suggesting to users the most relevant playlists 'matching their taste'.

> *In the past, we used to do this based on songs. We wanted people to share songs. Everyone would share songs, follow each other; see each other's songs on their feed etc. We tested this structure for about three months and noticed that the music feed is not like the Twitter feed, where you just scroll down to read Tweets. It was not an appropriate consumption experience for music. Now we see that instead of one song, there are playlists created by you based on your personal taste. It's your own selection. That's why right now we're trying to build the whole experience based on personal playlists.*
>
> *Therefore, the new product is about discovering personal playlists and discovering people through these playlists and then following these people and staying informed about them. The idea is that I should be able to listen to your Spotify or Deezer no matter which music service I am using. Similarly, someone who is not subscribed to any service should be able to listen to my playlists I created on Deezer. This is the main problem that we intend to solve at the moment* **(**Erdem 14/05/2014**)**.

In this model, partnerships between Cubic.fm and the streaming services are based on *data synchronization* between platforms. When a user connects her streaming service account to Cubic.fm, the system automatically pulls in previously created playlists on these platforms together with the behavioural history of the user. From that moment onwards, whichever action a user takes on Cubic.fm *syncs back* to the streaming service and vice-versa. Thus, the new structure enabled different streaming services to access each other's data through Cubic.fm.

The initial launch of this social media version was followed by another change concerning the structure of the user participation built on the platform. The main reason behind this decision was a recognition that the observed complexity of the system resulted in failing to engage users. A team member explains this as follows:

> *In our last meeting I told you that we were changing the structure, right. The reason was complexity. Our product has too many 'call to actions'. Let's say you go to Instagram, there you've got just one call to action: it tells you 'to share' your photo. You know you need to take a 'single action' to use this platform. But our product had too many actions. You go to Cubic.fm, it tells you "to connect" your streaming service accounts. You connect them. Ok, cool. Then it shows you a whole bunch of playlists by other people. But you had connected your own services to import your existing music, why do you see other people's playlists, right? But anyway, you've seen them, okay, good. Then there's the action of 'publish' for your own playlists, and 'follow' for other people's playlists. But then what is the main action of the product? It has to be through the 'follow' function. I mean, why does 'follow' work on Twitter? Because the only thing you do on Twitter is to tweet and read other people's tweets. So you either post something or follow. That's it. But in our product, you can both connect your Spotify account and publish your playlists. Do you transfer your playlist here so that others could listen to it, or do you just want to connect your Spotify and YouTube so you could see all of your music in one place? If it's the latter, why is there a call to action called 'publish'? You see, the structure was a bit complex and confusing* **(**Erdem 14/03/2015).

In order to overcome the observed *complexity problem*, the team decided to focus on finding one single 'call-to-action' to organize user participation. In doing so, they also wanted to make sure that this main call to action *effectively engaged users*. To this end, they decided to

build the new structure around *measuring the influence* of users. The idea was to develop a ranking system that facilitates user participation with a single and clear call to action: 'up-vote'. Similar to the 'like' button on Facebook or the 'thumbs-up' button on YouTube, the up-vote function was the main action organizing user participation and social interaction in the new system.

Fundamentally, this was a reputation system that ranked users according to their influence score. Calculation of the influence score was based on several factors. One of them was the publishing of playlists on Cubic.fm. Users could use up-vote the playlists shared by particular users to show their appreciation. Thus curators of the most appreciated playlists would become the most influential users of the platform and go higher up in the ranking list. Users had the power to affect rankings by their votes and by continually staying active on the platform. In the following vignette, one of the developers explains the calculation of the influence score:

> *In our database, we keep a user-tag score that signifies your Music ID. Before, we used to keep that score in the background. When one of the playlists became popular, it didn't have any visible advantage to me. There wasn't a social effect. Now we're trying to add the social effect of your playlists into your score. I mean, in the past, calculation of Music ID was more personal. Now it is a bit more social. We used to keep those data in the background without showing them to the users. We were implicitly showing it in the form of personalized recommendations. Now the user will see a score explicitly. This score will signify the "influence of a user's 'music taste' on others"* (Baris 26/12/2014, emphasis added).

The vignette above explains how the calculation of Music IDs changed from being purely a backend operation to one that is now partially visible to users in the frontend and affected by users' influence scores. In other words, the calculation of an influence score would now affect the ways in which music taste was calculated. As described in the previous section on the back-end operations of the platform structure, a user's Music ID was represented in the system quantitatively in the form of a user tag-score. This score was calculated through the measurement of the user's *interaction with music*. Incorporation of users' influence scores into the calculation of their music taste would add an additional variable into the formula: the measurement of a user's *social performance*. That is, the more users participate in the system by up-voting others' playlists or the more a user's playlists are up-voted by others, the more

their music taste gets affected. As a result, quantified versions of users' music tastes became even more dynamic than they previously were. Users could see how the influence of their music taste dynamically changes and evolves. This occurs in a similar way to the experience of the visualized Music IDs available to users in first version of the platform (See again Figure 5.10).

There was another crucial impetus behind designing user participation around a dynamic ranking system: *to encourage* users to create content and let them get *feedback* from each other. It other words, it was to let users see how others *react* to their curation in a notably more material way. This is accomplished through the counts of the upvotes that serve as the metrics underpinning the ranks. The expected result of this new structure was to increase user engagement. How well they perform in terms of curating good quality music content was directly reflected in a dynamic ranking system. Meanwhile, they were able to improve their rank by *continuously staying active* and adding more playlists to receive more reactions. Because the explicit appreciation from the other users affected their position within the community they serve for, users were expected to be more encouraged to produce content

> *Look at it this way: when you go to an Instagram account, you look at the number of followers users have. It is the same for Twitter… You may have a hundred playlists. But what matters is the number of people who listen to your playlist and how many people you influence there through the music you create. There is not such a structure anywhere. There's nothing about this. One exception might be the followers users got on SoundCloud. But we're trying to build a much more dynamic structure where users can react faster…Well, it will also have the following effect: when a user publishes a playlist on Cubic.fm she will not only create her own community (audience) but she will also receive reactions from her audiences very quickly. So when she receives good feedback immediately she will be more willing to do this over and over again* **(**Baris 26/12/2014).

In addition to the release of the up-voting system, another change the team decided to make was to create a distinction between the social and personal aspects of the platform. While the social aspects were being boosted with the newly introduced up-voting system, there was also going to be a section dedicated completely to the personal library of the users. This section was designed as an iTunes-like library, closed off to the view and reach of the other users. In this section, users could access all the music content they imported to Cubic.fm from the

streaming services. To further facilitate this social/personal division, the team developed a browser extension, which was going to function as a bookmarking tool to collect music while browsing the web.

The planned release date of this new version of the platform based on the up-voting system and the browser extension was early 2015. However, although the development was completed and the new system was already being tested in a beta version, this version was never released. The main reason behind the decision not to release these newly developed functionalities was the perceived complexity the system already had. To boost user participation in one way, Cubic.fm operated as a bookmarking tool, enabling users to centralize their music collections in one place by collecting previously dispersed ones to build personal libraries. On the other side, the playlist sharing and up-voting system was also designed to boost user engagement by improving the *feedback mechanism*, motivating users to create more after seeing their ability to influence others. With these two very distinct functionalities operating simultaneously, Cubic.fm was, in effect, trying to combine two different platforms with very different 'calls-to-action'. The team noticed that, in putting so much effort into increasing user engagement through various means, they failed to do it in a focused manner.

Considering that the departure point for the last release was to engage users with a simple and single 'call-to-action', the team decided not to release the new functionalities altogether. Instead, they chose to build the 'core functionality' of the platform around *personal use*. This way they could build a simple, easy to use platform focused on *engaging content creators* first. Given that the ultimate aim of the company was to centralize music consumption data in one place, enabling and motivating users to use the platform first as their personal library and centralize their music there became a priority for the team. Thus social features are put aside, but they remained in the team's roadmap to be implemented on top of a 'well-established core' at a later phase. As a result, the configuration of the platform was changed one more time, and Cubic.fm was launched in a private beta as a globally accessible online music library connecting different music sources in a single player.

5.4.3. Evolution of the Online Music Library Versions

The latest version of the Cubic.fm platform is a *personal music library* described to users as a *single access point* for all their online music. User can create an online music library on

Cubic.fm by 'collecting' songs as they come across them on the web, and they can listen to them on Cubic.fm independent of the source providing the music content. The Cubic.fm browser extension that functions as a *bookmarking tool* sits at the core of this new system. It also underpins the new vision of the company: *"to redefine the music library experience for the streaming era*." This vision is explained in detail by one of the co-founders in detail through an article shared on Medium.com as presented in Figure 5.39 below.

Figure 5.38 The article published on Medium.com in which one of the co-founders elaborate on their new vision. Source: "URL is the new MP3. Redefining music library for the streaming era." This is a web post written by one of the co-founders of Cubic.fm. It is published on Medium.com. Available at: https://medium.com/cubic-fm-stories/url-is-the-new-mp3-9d3182e3b063

According to team's new vision, the new motto of the company became the following: "*URL is the new MP3*."

> *Let's say you've discovered a Spotify playlist, you can simply add the URL, and that's it. So Cubic.fm is the place where you collect everything source-independently. Actually, we say: "URLs are the MP3s." In the past, we collected MP3s from everywhere. Now we're collecting 'URLs' because we are now living in the streaming era. When you send a song to a friend, you send its URL, not the song itself. This is the kind of vision we have now* **(**Erdem 16/12/2015).

The bookmarking functionality, combined with the functionality of connecting streaming service accounts in one place, seemed like a compatible form for the newly designed online music library service. The idea behind it was to build an iTunes-like library that connects personal libraries (user playlists) spread over different music services. But, unlike in iTunes, the Cubic.fm library would be composed of URLs instead of mp3s. However, the team noticed again that combining two functionalities could have led to further complexity. Therefore, they decided to stay focused on establishing the *core* of the platform around a *simple* functionality. The core of the platform is built on a 'save-to-cubic' action enabled by the browser extension. This browser extension renders 'connect service' functionality redundant, because the browser extension enables users to collect content by adding the URLs. Users can save the URL links of the playlists they created on streaming services. By doing so, they could keep all of their playlist all in one place. This way, the team ended up building the core of Cubic.fm platform with a single and simple functionality: enabling users to centralize their music collections in one place.

> *Before, we insisted on having users listen to music on Cubic.fm. Now we've seen that we cannot do that. You cannot tell the user to discover or listen to music at specific places. A user can listen to music anywhere she likes. What you need to tell users is that: they may listen to music wherever they like, on any service or blog, but what we're doing is to be with them (users) wherever they want to discover or listen to music, so that they could put it all together in one place* (Erdem 26/03/2015).

5.5. Conclusion

This chapter outlined the case study narrative. I first depicted the context of the investigated company's emergence, describing the background, vision. Thus I described the business model of Cubic.fm as a data-driven company aiming to centralize music consumption data to eliminate fragmentation in the digital music landscape. Afterwards, I portrayed the main unit of analysis of this research project: that is, Cubic.fm's product-platform. I described in detail the three different forms this platform design has taken. Here I described the basic structure and the backend operations of these three different versions of the Cubic.fm platform, namely as a: 1) crowd-sourced radio service 2) social media platform for music discovery, and 3) consolidating personal music library. Following this part, I outlined the processes through which the platform was reconfigured to take on these three distinct forms. Here I described the reasons behind the changes made in platform system, as explicated by developers in

interviews. In doing so, I identified the problems with and the challenges for the existing system that were highlighted by the Cubic.fm team, while pointing out the perceived benefits of the newly designed system. In the next chapter, I will draw on this case study narrative and present the main findings of the study.

6. Findings

6.1. Introduction

The previous chapter outlined the empirical case narrative by giving an overview of the investigated company, the design of different versions of the platform, and the process through which the platform had been developed and evolved. In this chapter, I present the key findings of this case study. These revolve around the challenges faced during the development and optimisation of the Cubic.fm platform, and the ways in which they had been addressed to result in the reconfiguration of the platform three times. First, I discuss the key challenges and concerns that the Cubic.fm team had while building the platform. These paved the way for the changes in the design and the related data-based operations of the platform. Then, I move on to elaborate on the strategies, tools, and techniques used by the team to address these challenges. I conclude the chapter by summarising the findings and then proceed to analysis and discussion.

6.2. Major Challenges and Concerns

The challenges faced during the development of Cubic.fm as a social media platform for music discovery are subsumed and discussed under four prominent themes. These are: 1) ordering music; 2) data integration and standardization; 3) establishing and adjusting the core interaction; 4) increasing user engagement. The Cubic.fm team primarily addressed these challenges using analytics and the accompanying measurement techniques. These are used to track and analyse user behaviour in ways that help to tweak the platform system in order to create the type of user behaviours that are desired and expected for the scale and success of the platform.

6.2.1. Ordering Music

The empirical narrative shows that the evolution of the platform from an activity/mood-genre-based crowd-sourced radio service to a social media platform was mainly due to the perceived power of ‘user-generated’ *data* and *content*. In addition to being very good indicators of users’ music taste, user playlists are also considered to be powerful sources for organizing, grouping and making sense of music in ways that do not necessarily have to fall within the confines of traditional genre categories. User playlists are found to be highly

efficient means for analyzing and ordering music in the most meaningful, reliable, and scalable way. Reliance on user-led categorization simplified the computation of relevance between users' music taste and the playlists. Thus it helped the platform to better perform its *matchmaking* function.

In its crowd-sourced radio version, Cubic.fm was built on *algorithmic* curation, that is, the *automatic* identification and ordering of music objects using machine-learning algorithms. To be able to do so, the team developed a taxonomy system to train their algorithm. This taxonomy was based on their 3-4 years of experience in music retrieval techniques and recommender systems, which requires extensive knowledge on sorting music genres and the transition between them. That taxonomy was developed mainly due to problems in finding reliable metadata for music. Its primary function was to ensure that the initial set of songs used in training the machine-learning algorithm aptly fit into the activity/mood-genre category of the radio stations (algorithmically curated playlists). In other words, this taxonomy was developed so that the algorithm could verify the appropriateness of newly added songs reliably. New songs were added to radio stations either automatically through web crawling or through the contributions of users as they 'cubed' songs to stations. Consequently, identifying and sorting music content in the first version of Cubic.fm was reliant on machine learning algorithms guided by an editorial taxonomy.

Over time, the Cubic.fm team observed that most listeners already had personal playlists they created on other music services. In the team's view, those playlists are very good sources to be imported onto Cubic.fm as already available content which has been generated, clustered, and sorted by users according to a particular logic. They are considered to be very rich data sources, not only for understanding listeners' music tastes, but also for sorting out music objects. Yet they remained idle on other services. With these factors considered, user playlists seemed like a more effective solution for ordering music compared to the editorial and algorithmic work the team had to do in order to create activity-genre based radio stations in their crowd-sourced radio station version. The following comments made by Cubic.fm team members highlight some of the concerns behind the decision to move towards user generated playlists:

Right now, identifying and categorizing music is passed over to the users **(**Baris 05/06/2014).

> *Moving towards a user playlist-based system is much more scalable because we no longer have to do other type of radio stations. I mean, we don't need to test or validate things in the way we used to do algorithmically in the previous system. So the user-generated model is much more scalable* **(**Erdem 05/06/2014).

The same logic is also operationalized in the following personal library version. The standalone form the platform has taken in its latest version focuses on personal use rather than social features. This was done mainly to solve the 'chicken-egg problem' experienced by most platforms in initial phases. In order to trigger network effects, the platform needs both sides (both producers and consumers of content) to be engaged. But, due to the observed lack of engagement for content production, the team decided to focus on attracting content 'producers' first. To this end, rather than being asked to connect streaming services accounts to import their existing playlists, users are given a tool - in the form of a browser extension - that helps them to collect the songs they come across while browsing the web. Once saved in Cubic.fm, songs can be grouped, arranged, and categorized. That is, they can be ordered and sorted out on the platform. In this way, although they are referred to as 'channels' in the latest version, these dynamic playlists that have been created and curated by users provide the Cubic.fm platform similar means of ordering music as the user-generated playlists of its former version.

6.2.2. Establishing and Adjusting the Core Interaction

Design of the core interaction plays a crucial role in organizing user-platform participation. The Cubic.fm team, in search of the most engaging core interaction for the social media platform they were building, have experimented with the platform's design several times. While doing so, using analytics, they have continuously observed and measured the ways in which users react to changes made in the platform system. Based on these users reactions, they have tweaked the platform system numerous times or released new sets of features to address observed problems in the interaction between users and the platform. These features are then tested again using the same means. Thus the team has relied on a cyclical process of measuring user behaviour and adjusting platform functionalities. This process sometimes ended up with a complete shift in strategy known as 'pivoting'[5]. In the following vignette,

[5] The term 'pivot', when used in the start-up context, describes the path of finding the right customer, value proposition, and positioning. According to The Financial Times Lexicon, what pivoting implies for a company can be described as "keeping one foot firmly in place as you shift the other in a new

one of the developers explains the importance of finding the right set of functionalities to address the needs of the users and establishing them as the core features of the platform:

> *The first question is what cubic.fm is, and, what kind of problem it solves… Then we decide on the features of the product (platform) that will solve this problem. Then, let's say, a user comes and logs-in and becomes a user. Whether he/she uses these features or not after this initial traffic will become a criterion to understand if we are able to solve the problem we aim to address. To do this, first I define a user as 'active' or 'engaged' user and say that she will use certain features. Now, I need to check to see if these features are my product's main features or not. Are these features really the ones that solve the problem I want to solve? If I don't make sure that these are the best features, then I start to lead the product to the wrong direction. I begin with an error* **(**Baris 27/06/2014**)**.

Throughout the Cubic.fm platform's evolution, the vision of the team regarding music discovery has always been a social one. This is evident since its initial crowd-sourced radio version. Nevertheless, in the process of optimizing the platform's core interaction, the team also had to solve the chicken-and-egg dilemma, which led to sacrificing social features for the sake of increasing user engagement and content production. The last version was launched as a 'stand alone' product focusing on establishing the core by engaging the one side of the platform first, through a personal library without any social features. However, integrating social interaction as a peripheral feature remained as a future goal on the roadmap of the team. Thus, social features are set as 'peripheral' elements that will be built on top of a 'core' designed primarily around content production.

The initial challenge that the team experienced with regard to the design of the platform's 'core interaction' (user-platform participation) was faced in its crowd-sourced radio version. It was related to the special nature of music and the ways in which people interact with it. In the first version of the platform, user-platform participation was built on a feed structure in which individual songs were the 'value units' of user participation. Designing the core interaction in this way did not create the expected user engagement. That is because music is inherently different than other types of media, and people interact with music differently than

---

direction." Available at: http://lexicon.ft.com/Term?term=pivot In this regard, we saw that Cubic.fm pivoted three times.

they interact with text or image-based content. As observed and reported by the Cubic.fm team: 1) it takes at least a few minutes to listen to a song, while people can read tweets in seconds; 2) people may enjoy listening to songs repeatedly, as opposed to reading a tweet or seeing a photo just once; 3) song transitions matters to listeners as much as the songs themselves. For these reasons, the team decided to abandon the feed structure and redesign Cubic.fm's core interaction.

In the second version of Cubic.fm, user-platform participation was organized around playlist sharing. In this way, the value unit of the core interaction became user-generated playlists. Facilitating user participation through user playlists was considered as more scalable design choices. However, in taking the form of a social media platform for music discovery, the second version of Cubic.fm faced new challenges from the outset. In its previous crowd-sourced radio service version, the majority of the content consumed on the platform constituted of editorially created, algorithmically developed playlists. As Cubic.fm adopted the social media model in its second version, the production of music content and the ordering of music were outsourced to users. So, unlike its predecessor, it was completely reliant on user-generated content. This was the main difference between the two versions and, with this change, user participation became the most essential concern:

> *Like I said, we wanted people to come and create playlists, tag them, write a description, put a nice photo up there. We didn't want them to just listen to music and leave, but to contribute content as well* **(**Ozgur 18/12/2015).

As a social media platform, Cubic.fm needed to motivate users to continuously participate in creating and tagging playlists. However, observation showed that user behavior leaned towards consumption; that is, the majority of the users tended just to listen to music available on the platform without creating playlists themselves. 'Consumers' outnumbering 'producers' on a discovery platform is perceived to be a highly concerning drawback for the future of the platform. Active content producers are able to attract consumers who just listen to music on to the platform. However, 'mere consumers' without the 'active producers' creating playlists for them had no value for Cubic.fm platform in its initial stage. Having observed that the number of producers was not enough to feed the consumers, the team decided to focus on tackling this problem.

The developers of Cubic.fm first noticed that the complexity of the platform was discouraging producers. In order to overcome this problem, the team decided to simplify user participation. Moreover, they decided to explicitly gamify the core interaction by establishing it on top of a ranking/reputation system. The idea was to motivate and encourage playlist producers through a dynamic feedback mechanism. Thus, the team developed the 'up-voting system'. The structure of user participation in the up-voting system was more or less the same as Cubic.fm's core interaction, as analysed previously. However, there was a new call to action: the 'up-vote'. The 'up-vote' button resembles the 'like' button on Facebook or the 'thumbs-up' button on YouTube. The difference lies in the ways in which the system dynamically 'ranks' users based on the 'up-votes' they get. In this version, the system calculates an 'influence score' for playlist producers. Their score changes based on the up-votes they get every time they create and publish a new playlist. This affects their rank in the list of *influential/popular users*. As a result, producers are expected to be further motivated and continue to create playlists:

> *Up-voting was designed for our core users. These were the people who would make an effort to create playlists, get motivated by the feedback coming from the platform, and keep on creating more playlists* **(**Ozgur 18/12/2015).

In parallel to the up-voting system, the Cubic.fm team also developed a 'browser extension' that functions like a bookmarking tool as it enables users to collect music across the web. The team planned to release this extension together with a 'widget' enabling the circulation of the Cubic.fm playlists on the web. These two tools were designed to work in a complementary fashion to each other in order to increase engagement and growth. However, this version that encompassed several different functionalities was not released due to the complexity added to the platform's core interaction. The team decided instead to focus solely on designing the 'core'. Given they had also been trying to tackle the problem of engaging content producers, they decided to establish the platform 'core' based on *personal use*. Thus, social interaction is set as a 'peripheral' feature that will be integrated as one of the add-on functionalities in the future:

> *We should develop such features that would really help people to create playlists. But crucially they should use these features for their own needs; they should create playlists for their own libraries so that there will be enough content on the platform.*

This way at a later stage we can welcome those users who do not necessarily create playlists but would like to use Cubic.fm just to listen to music (Ozgur 18/12/2015).

So, in latest version of Cubic.fm, the team focused on developing the 'producer side' of the 'core interaction' first. The 'value unit' is changed one more time to become individual music items (i.e., URL link of a song or a video). As an online music library, the latest version of Cubic.fm enables users to save the songs they come across on the web, keep them all at one place, group them in different playlists, and categorize them by tagging. In this latest version, playlist producers are the only consumers of the playlists they create. They can play the songs they collect from different sources on the web via the source-independent player provided by Cubic.fm. The following passages summarise the reasons behind this change:

We should have people use this platform like a library. We should completely remove the 'discovery' feature for now. We could add it back again later. But when we put it on day one, we change our target population a bit. We don't want these guys to be inside the product on day one, okay. Because, they are really not the kind of guys that have a value add for us. We need a clear vale proposition, and from day one onwards we need to get to guys who can add value to the system. If these guys create lots of content, we will already be at a point where we could offer the mass a lot of content to discover

(…) Because, look, when you offer ready-to-consume content at first, like I said, these guys come over immediately thinking, "I'll just listen to music there." For now, we do not want these guys to be there at all for a while. It may sound wrong, but it's not. Actually, almost all of the prominent social media platforms started like this. We discussed that several times among ourselves too. We couldn't name a single social media platform that didn't start this way. They really start out as closed communities and then they grow big. We should be able to do this… Nothing really changes in the background; the number of functionalities may again increase in the long run. But first of all, Cubic.fm needs to be able to propose you a value, even if there is not even one more person on the platform. You need to have a reason to come here. We just want to give you that reason **(**Erdem 26/03/2015).

6.2.3. Data Integration and Standardization

Another core challenge the Cubic.fm team faced is the complexity of the data infrastructure upon which the platform is configured. Building a data infrastructure for centralizing data aggregation on music consumption involves two key challenges: 1) connecting too many APIs and 2) dependencies on other services.

i. The Challenge of Connecting Too Many APIs

Throughout the evolution of the Cubic.fm platform, no matter how many times the platform configuration has changed, the main objective of the company has remained the same: to create a platform that centralizes data on music consumption. The aim is to eliminate the fragmentation of data and user experience among different services by connecting their users on Cubic.fm. The developers tried to accomplish this goal in several ways. For instance, in the second version this was attempted by locating Cubic.fm as a platform in the middle of different music services. To be able to do so, the platform enabled users to import their playlist from other services on to Cubic.fm. It also asked users to sign-up to Cubic.fm using their social media services to aggregate an initial set of personal data in order to overcome the cold-start problem. Consequently, Cubic.fm's data infrastructure was composed of data aggregated using several different sources, such as: social media services, streaming platforms, metadata services, weather applications etc. This was fundamentally accomplished by using the APIs (Application Programming Interface) of these services. The following figure shows how one of the co-founders explains this API dependency while answering questions from wider public, in the launch of the latest version through the Product Hunt Platform

Figure 6.1: A screenshot from the launch of Cubic.fm in the ProductHunt platform, where one of the co-founders explains how they use APIs of various streaming services to stream music free of charge. Retrieved from the ProductHunt platform, 'Cubic.fm'. Available at: https://www.producthunt.com/posts/cubic-fm-2

In the following vignette, one of the developers explains the difficulties, as well as the opportunities, involved in this process:

> *When we accomplish this, then we will become a separate platform on our own. I mean, I explain all of it so easily, but integrating that many services isn't easy, really. It's extremely difficult to get this flow right. What's even more difficult is to keep the data integrated in the database through these services. We import playlists from Deezer, Spotify etc. We should always be able to match the songs in those playlists with other services. We really worked for a long time to accomplish this. It's not easy. But we did build a service that works well for now. Cubic.fm will be a great platform for creating next generation services too. Because, instead of developing applications by using the infrastructure of each service separately, these new services will have the chance to develop applications that are integrated with all other services using Cubic.fm's infrastructure only* **(**Ozgur 19/04/2014**)**.

One of the most important challenges that the team faced while building its data infrastructure was the integration of YouTube. The main challenge with the YouTube

integration was related to the user-generated nature of the service. Given that the core aim of Cubic.fm was to centralize and standardize music consumption data and music metadata, the company was keeping records of the different ID's of the same song, as found in different music services. In this way, the service was able to find the same song in different services by matching its ID's. Since it could stream the music from different sources by matching their ID's, it enabled users to access the playlists of the users of different music services

It was easy to find the ID of a particular song in streaming services because they provide *structured* metadata. Therefore, it was easy for Cubic.fm to find the same songs on other services, retrieve their ID's, and match them in its own database. However, doing so was not easy when it came to YouTube. YouTube content did not provide any standardized metadata in the structured way streaming services did. So the Cubic.fm service had to work on string (text) matching to integrate YouTube into its infrastructure. The same was also the case for SoundCloud, another important user-generated content service for music. In the following vignette, one of the Cubic.fm team members explains the difficulties related to integrating user-generated services into their data infrastructure:

> *This necessity arises due to the fact that we already have a song database, and we already know the IDs of songs on Spotify, Deezer, and Rdio; but in order to play them for free, we also need to know their URLs on YouTube. For example, "4 minutes" by Madonna. We know how to find it on Deezer but not on YouTube, and we need to find that. And the only way to do that is through the title that the user enters at the time of uploading. Therefore, on YouTube we have to match title of each song one by one. Of course, we do this with script, not by hand. Also, next to the song title, there can be additional information in parenthesis, such as "original mix," "live," "mastered version," or, it could be the year of the song. There are thousands of cases like this* **(**Erdem 14/08/2014).

Integrating these services into its data infrastructure enabled Cubic.fm to categorize the unstructured data generated on those services. Thus it started to build a data infrastructure composed of standardized and matched ID's of the music content and music metadata available on different music services.

ii. Challenges Related to Dependencies on Other Services

The goal to connect different services in one platform resulted in a data infrastructure built on a complex network of APIs. This structure poses several risks for the platform due to its dependency on other platforms. Using APIs requires Cubic.fm to dynamically respond to all sorts of changes made in any of those APIs in order to avoid ruining the flow of its operations. Likewise, any action Cubic.fm wants to take is also dependent on the constraints generated by those dependencies. For instance, during the integration of these streaming services, the team needed to wait for Spotify's web API to be released due to the limitations that the desktop API put on Cubic.fm's operations. Thus the launch of the platform had to be postponed until Spotify released its web API. Similarly, the team needed to complete the integration of other services to be able to launch the browser extension in order for it to function properly. In the following vignette, one of the Cubic.fm team members reflects on the difficulty in building a data infrastructure that is dependent on the contingencies of other services:

> *Oh, and, by the way, towards the end of the summer, Soundcloud started to set limits for its API…I mean we've always got some load on our shoulders. All these services, APIs etc., at each step, we have the burden of working in a way that makes us dependent on this stuff… At each step we take, we depend on other services' terms and conditions. And you see these terms can suddenly change like that* **(**Erdem 16/12/2015).

6.2.4. Increasing User Engagement

In order to design and develop an effective social media platform for music discovery, the Cubic.fm team had to tackle several challenges. Most importantly, they needed to *motivate* content *producers to create* playlists and *consumers to react* to the produced playlists. It was only through balancing out the production and consumption on the platform that Cubic.fm could create the necessary conditions for data production that satisfies its needs as a data-driven company.

Analysis of the evolution of Cubic.fm platform with respect to 'the means of content production' demonstrates that there had been a drastic change in the operations and focus of the company following the shift from an 'algorithmic curation' of content to a 'user-generated' social media model. After being launched as a social media platform, Cubic.fm

needed *continuous* user participation and engagement. The following vignette illustrates these points:

> *We could have done this curation (creation of playlists) ourselves too. And actually we were doing it in the previous version. But it is not scalable. Therefore, somehow, we need to bring together the users who consume and the users who produce. We need to bring them together in a correct ecosystem. The reason why we're using these analytics is to make this ecosystem work better. What we're doing right now is not to make some incredible amount of sense out of this aggregated data, but, rather, trying to create a system that makes you want to be a part of that ecosystem…But I don't want you to see us trying to build crazy data mining or recommendation models once we've built this. We are, rather, interested in making this platform more connected. Therefore we are looking for ways to make more people part of it. This is what adds value to data-driven businesses.*
>
> **…***Even if you write really good algorithms, as long as I have a bigger and more connected sample, mine will surpass yours… This is what we came to understand in the process. We first began with the algorithms but later realized that it's more important to have more data at hand and also for that data to be more connected and unified. So, in such an ecosystem, the size of the sample matters much more than the algorithm. That's why now we're allocating our time quite differently compared to the initial set-up. Before, our focus was more on improving the algorithm, but now we're trying to think about questions such as 'what is the best infrastructure and configuration to empower this set-up?' or 'what are the channels through which a user interacts with music on the web and how can we put these channels together?' etc.* **(**Baris 14/08/2014).

To be able to build an ecosystem, as described in the vignette above, the most immediate concern of the team became securing continuous content production. The team refers to content producers as 'curators'. These are the users who stay active on the platform, produce and categorize playlists, and then become further motivated by the reactions they receive from other users (consumers) for their curations. Curators are regarded as the engine of the platform's scalability and growth. Therefore, as the structure of platform shifted towards a social media configuration, the core objective of the team is accordingly shifted from improving its algorithms to engaging 'curators', as explained by a team member below:

After we switched to the user-generated content, our core users become the curators. Because, if we were resolving this, let's say, with the algorithm, we would focus on how we could make the algorithm better, how we could improve it. But we believe that it's more valuable to use the playlists that millions of people are already creating instead of improving listening experience by the effort we put on the algorithm. I mean, that's how Pinterest and Twitter came about. There were millions of people keeping blogs, taking photos. These services were distinguished because they were able to gather all of this stuff at one place. If we can do the same thing, we could build a strong discovery system on top of that. The most scalable set-up is this one. I mean, the effort that goes into it makes more sense than developing the algorithm and the result we would get from it **(**Baris 27/06/2014).

Cubic.fm's focus shifted towards engaging curators because, as a social media platform relying on user-generated content, it needs its users to continuously create new playlists. Users also need to sort the content they produce by 'tagging' their playlists. In this way, playlists could be rendered *identifiable*, *findable*, and, hence *matchable*. As a result, it becomes possible for Cubic.fm system to perform matchmaking to facilitate music discovery. For these reasons, the platform needed to find ways to engage the 'active' playlist 'producers'.

i. Engaging the 'Active User'

There are several levels of user engagement. These levels are referred to as 'user states'. The Cubic.fm team heavily relies on analytics tools to identify and analyse different user states. These are deployed to make sense of user behaviours as well as to identify and sort out the characteristics of different forms of engagement:

So you can say, "If the user uses my product once a month, it indicates a state of activeness to me." Or maybe it's not like this for the music product. It could be the case for LinkedIn, but maybe for a music product you need to use it every day. It is up to us. We define this. We define whatever type of behaviour we want in a particular user state. Or we associate a particular state with a particular type of behaviour. Then we ask which users are in this state. To understand this, we track their behaviours. Then we go: "Hmmm, what are the common characteristics of these users? Did they add a lot of songs? Do they have lots of play counts? Or how many

songs did they add in the first few weeks and so they became active users?" I mean, we are trying to understand the characteristics of this active state so that we could move the product in that direction **(**Erdem 16/12/2015).

The 'active state' described above is often defined with respect to its opposite; that is, 'passive' user behaviours. For instance, when the team observed that user behaviour was leaning towards mere listening - in other words, when consumption outnumbered creation activities - concerns were raised about the scalability of the platform. Thus, the most immediate and urgent goal of the company became engaging users in ways that can convert them into 'active' and 'retained' producers.

In order to motivate content production, the team redesigned the platform several times. They experimented with different configurations of core interactions to organize user platform participation in the most engaging way. Finally, to overcome the chicken-egg problem, they decided to organize user participation by prioritising engaging producers and solely focusing on facilitating content production. Despite the differences, the rationale behind each configuration was more or less the same. The two vignettes below illustrate this point:

In the playlist-sharing version, we were trying to build a system that, at the end of the day, would work by itself. Certain people come and create playlists and categorize them. That way, other people can make searches with whatever tags they want and access playlists. At this point, if we build this balance and make the system work, we do not have to manually interfere. We will already have the necessary number of people creating playlists, and the people are able to access these playlists by these tags…The core users we need for this set-up were determined as the kind of people who would make the effort to create a playlist. A person who would create the playlist and then get motivated by the feedback that comes from our platform and keep on creating playlists… We are trying to reach those people, but the playlist-sharing set-up didn't help us reach out the target users that we have been looking for **(**Ozgur 18/12/2015).

Our ultimate goal is data aggregation. Because if we're going to be a big company one day, this is going to happen thanks to the data we will have gathered. What we need for it is to: 1) reach more people, 2) make users create data on Cubic.fm and be able to move their existing data in other services to our platform. This goal has not

changed at all. But right now, in order to make that possible, we just need to be able to reach the right set of people…I mean, our goal is still the same. But to make it possible, we need to 'lock' curators in here. We need people to spend more time here. They need to gather songs here. But they need to be doing this for themselves, so that, when you make this content discoverable, you would be able to bring-in the entire mass **(**Erdem 26/03/2015**).**

Changes made in the platform's design have also altered the definition of the 'active' and 'engaged' user for Cubic.fm. In other words, the type of expected user behaviours has altered in accordance with the new features of the core interaction that organizes user platform participation. For instance, in the second version of the Cubic.fm platform, 'active' and 'engaged' producers were defined as users who would 'create', 'publish' and 'tag' playlists. In the subsequent stand-alone personal library version, users defined as 'active' were the ones who continuously added songs found online to enlarge their personal Cubic.fm libraries. The structured procedure of 'creating' a playlist used in the former version, as composed of the four mandatory steps delineated in the *Road-to-the-Perfect-Playlist*, is significantly different than forming playlists in the latter by 'saving' songs through a browser extension whilst browsing the web. So, in this regard, attributes used in the definition of the 'active' and engaged' user changed according to the features of each system. This illustrates the extent to which the design of core interaction influences user-platform participation. What is designed to engage users thus results in shaping what users are expected to do. This indicates that the design choices are prone to create a *self-fulfilling prophecy* effect.

By deconstructing assumptions underpinning the idea of "engaging 'active' users", one can observe how the Cubic.fm team has a particular type of 'target' user in mind. These users are 'already out there' and 'already active'. In this regard, creating engagement is seen as just a matter of motivating the 'right' user for participation. In other words, 'if only' the team could develop the right set of features and functionalities, 'active users out there' could engage with the platform. A team member explains this as follows:

We're changing the target group a bit. Instead of mass users, we're shifting the focus to the people who actively discover music and trying to get them engaged […] We need to get these types of people because they can take us to other passive people and they are indeed taking us to other people. But the passive ones don't take you to anyone (Erdem 26/03/2015).

In the following vignette, a developer reflects on why they failed to engage 'the type of users they needed' in different configurations of the platform, and what they decided to do after they observed this. This reflection discloses the rationale behind the design and development of the design of the final, stand-alone version of Cubic.fm platform. In doing so, it not only illustrates the latest strategy used for engaging users in the production of data and content, but it also summarises well how the platform design is optimised in ways that can shape user behaviour to produce the desired outcomes.

> *I mean, if we look at it from a high level, we wanted to appeal to the people who don't just want to listen to music, but are also 'willing' to create it. That's why we developed a product just for them in the last version (…) We have realized that the number of people creating playlists on Cubic.fm wasn't enough. So we decided to develop a specific tool for them. Our approach became developing features that would really help those people create playlists. The idea is that they should use them for their own needs, and, as a result of this, they would create playlists. Thus, once we reach a sufficient number of playlists, those who prefer to just listen to music can find playlists inside* **(**Ozgur 18/12/2015).

Another example that can further illustrate this point is the standardization of the playlist creation process. Referred to as the *Road-to-Perfect-Playlist*, the playlist creation process in the second version was composed of several mandatory steps. This was a highly structured process, pushing users to create content for the platform in a *particular way*. Playlists that failed to fulfil the requirements specified in this process remained as draft playlists with a progress bar showing the completion rate of the playlist. The main function of this progress bar was to inform the user about the remaining work needed to complete the playlist and render it eligible for sharing with other users. The following vignette illustrates the reason behind standardizing the content production process this way:

> *We found a way to ensure that every playlist is tagged in the system. The logic is the following: we have two type of playlists available in the system: draft and published playlists. Any playlist that's been created on Cubic.fm is a draft playlist and if the user wants to publish it, he/she has to fulfil certain criteria. What are they? It needs to be tagged; it needs to have an image visualizing its mood and context, and it needs to have a certain number of songs, let's say 10, but we haven't decided on the number*

yet… Once all of these have been done, then it becomes a publishable playlist for us, and can be published

…This provides several benefits. For instance, we had the problem of not having unified data. We were able to see the tags in some playlists, but we also had playlists with no tags at all. When the user interacted with any of those playlists, he/she wasn't sure what kind of music he/she was going to find in them. That's why any playlist created in Cubic.fm is now a draft playlist and it can be published only if it fulfils the criterion we set. Similarly, any playlist that you bring from your music streaming service is also a draft playlist, and we say that you can only publish it after completing the necessary steps **(**Onur 18/07/2014).

As this example illustrates, the ways in which the design of a platform is structured and the user platform participation is organized are significantly linked to the *data* and *content* production concerns of the developers. The vignette above (Interview 40) illustrates how the platform is designed in a particular way to ensure that the content generated by the user is also tagged properly. Standardizing playlist creation in this way eases the process through which content is produced and categorized by users. However, standardization alone is not enough to create the ultimately desired outcome; that is., *continuous* user activity around data and content production. It needed to be complemented with other strategies.

Standardized and stylized user actions help to structuring user-platform participation and push users towards particular behaviours *once* users initiate an activity. However, users cannot be pushed to initiate activities, since participation on social media needs to be 'voluntary'. In other words, a user can be forced to do things in particular ways, but they cannot be pushed to initiate an action. They can only be *motivated*, so that they can do required activities 'willingly'. In this regard, motivating users first to initiate an activity was also the part of the challenge to engage active producers. In the next section, I analyse how users are motivated to participate in data and content production willingly, and discuss the implications of the techniques used to increase user engagement.

ii. Organizing User Participation to Create Reactivity

In order to motivate users to participate in data and content production willingly, the Cubic.fm team experimented with several techniques. One of the incentivising techniques used is game mechanics. As pointed out in the literature, providing users feedback in order to

trigger intrinsic motivation is the most common gamification strategy used on social media. The basic aim behind implementing game mechanics into a process is often to create 'reactivity'. It can be observed that the Cubic.fm platform is designed in ways that purposefully invite reactivity, given reactivity is defined as: making people change their behaviour in reaction to being observed, measured and assessed. As explained by one of the team members below, one way of creating reactivity is to create a 'willingness' to initiate and complete particular actions. Rather than explicitly pushing users towards a particular action, initiating it on behalf of them or helping them complete it is, to a certain extent, assumed and expected to result in higher completion rates. An example of this is the case of the automatic tag suggestion feature embedded in the playlist creation process:

> *This week we have moved to a new set-up in playlist creation. First, we look at how many playlists users got and how many percentages they've completed. We want to make users complete it step by step. We analyse the playlists and suggest some tags. We say, "We have added some tags to get you started. Add tags to help people find your playlist." So, we analyse his/her playlist through its scientific metadata and suggest tags to ease the process for her. For example, this is a "chill-out" playlist and it got the tags of "easy-listening," "launch," "mellow," so… Or this one got "electronic," "launch," "loud." This helps the user quite a lot. I mean, users could still change the tags we suggest to them if they want to. These are just suggestions to help them. It's a bit like the Facebook login - because I've already started a step for you, you are more likely to complete that step. This is how we began to teach tagging playlists* **(**Baris 18/07/2014).

As this vignette reveals, user-platform participation is organized in particular ways to create *activity* through *reactivity*. These can be traced in most of the choices underpinning the design of Cubic.fm. Examples are wide ranging and include: presenting users with metrics related to user activity on playlists; starting and easing a process on behalf of users (as in the case of auto-tagging); ranking users based on their influence score (as in the case of the up-voting system); showing users how well they perform with a progress bar and informing them about the remaining amount of work required to complete the activity (as in the case of playlist creation process); alerting users about the reactions the content they generated receives through a notification system; or showing users their visualised Music ID's as indicators of their 'mastery levels' on certain music genres etc. All of these features are

purposefully designed to motivate users to initiate or complete an activity 'willingly' by reacting to a certain form of assessment or performance evaluation. In other words, it can be observed in all of these examples that the techniques of quantification, assessment, and measurement are purposefully employed to create reactivity.

iii. Tracking User Behaviour to Create Reactivity

Stylized social media actions used in the encoding of user-platform participation create standardized and measureable activity types that help in the tracking of user behaviour through analytics. Analytics tools play a crucial role in the optimisation of the design to create desired user behaviours. They are used to monitor different aspects of user behaviour (e.g., the completion time and rate of an action, the termination of a task at a particular stage, etc.). They also make the measurement of user 'reactions' possible. Moreover, as specific means used to measure online user behaviour, analytics help with assessing the performance of the design features in their function to engage users. There are two important points concerning the role analytics play in this process of: 1) spotting problems and 2) measuring the reactions users give to new features that are implemented to solve the spotted problems. In the next part (Section 8.3), I elaborate on these points and analyse how analytics tools and corresponding measurement techniques are used in Cubic.fm's efforts to address the challenges and concerns analysed in section 8.2.

6.3. Tools and Techniques Used to Address the Challenges

The Cubic.fm team relies on particular tools and techniques in response to the challenges described in the previous section. One prominent case in point is employment of analytics to continuously measure user behaviour. Decisions that led to changes in the configuration and, hence, the organization of user-platform participation, were often made after a process of analysing user behaviour using analytics tools.

The key issue in measuring user behaviour using analytics tools is to continuously monitor in real time whether users are taking the desired and expected actions on the platform. The term 'conversion' refers to occurrence of a desired action. In other words conversion happens when a user takes an action that the platform owners/developers 'want users to take.' Accordingly 'conversion rates' are the percentage of the number of conversions calculated by dividing the total number of conversions (desired actions) by the number of visitors. This

process is known as 'Conversion Rate Optimisation' (CRO), and analytics programs are the key tools used to increase conversion rates.

Conversions are defined by each platform's unique objectives and vary depending on the type and objective of the platform. An online guide[6] to conversion rate optimisation defines the relation between 'conversion', 'user behaviour' and 'measurement' as follows: "*Whatever it is you want your visitors to do, this action is what you are going to measure and what you are looking to optimize."* In this section, I draw on the Cubic.fm case study narrative to shed light on this relation between the notions of 'conversion', 'user behaviour' and 'measurement' by examining how a social media platform continuously uses analytics to measure user behaviour, and, in turn, use this data for conversion rate optimisation to create the desired user behaviours.

6.3.1. Measuring User Behaviour with Analytics

The main tools used by the team were Google Analytics and KissMetrics. These are used to make sense of user behaviour through the analysis of the data aggregated on the platform, as described in Section 8.2.2. These analytics tools are used to measure both micro- and macro-level interactions. Google Analytics, as one of the team members puts it, is used to measure the platform's relation to its external world. KissMetrics, on the other hand, has been the key tool used for understanding the dynamics of the 'micro-interactions' between the platform and its users. Moreover, while Google Analytics is used to gain aggregate-level insights about user behaviour, KissMetrics is a tool for measuring user behaviour at an individual level. Aggregate data enables developers to identify the most effective user acquisition channels or to monitor growth rates. When used in combination with individual-level data, such aggregate-level data also help the team to measure user engagement and fine-tune the elements of core interaction organizing user platform participation. A team member explains this as follows:

> *I can see everything on Google analytics, like how many people came through Twitter, for example. Let's say 100 people came through Twitter; I can also see that 30% of them have signed up as well. In order to do that, I need to customize Google Analytics for myself. I define an event, saying the action of signing up is this or that,*

[6] "The Beginner's Guide to Conversion Rate Optimization" -Chapter 1: What is Conversion Rate Optimization. Available at: https://qualaroo.com/beginners-guide-to-cro/what-is-conversion-rate-optimization/ Last retrieved 26 August 2017.

and then I tell the system to send me a notification for the users who have taken this action. This way, I define it as a 'conversion'…For example, between the dates I chose, there were 932 sessions by 700 users. This means that some of these users came and listened to music at different times of the day. This one here is the number of users, and this is the session of the user. The user displayed this many pages on average, and a session took this long approximately. So this number of users came and stayed listening to music for such an amount of time on average that many of them left after the home page.

…Here, we can see the characteristics of people, like where they came from, which device and browser they're using, how many of them are on a mobile device and how many on the web… For example, these, who came, have all been converted. There is no one that bounced. They have taken action, displayed three pages, and stayed for three minutes on average. They didn't listen to music for a long time; they just came to see what this is about. Oh, look, one of these two even got registered. We are the ones who define all these 'conversion events'. For example, 'register after landing' is an event as such. You define a goal like this, and it tells you how much of it has been accomplished. I'll explain it to you with an example. The number of people who have searched 'Cubic.fm' on Google is 26… Of these 26 people, 16 came for the first time. 30% of the people who came this way took no action, not even clicked on the 'join' button. The rest displayed five pages on average, stayed 22 minutes, listened to music and 8% of them got registered. Two of them came by searching 'Cubic.fm' and signed up through Deezer or Rdio **(**Erdem 11/07/2014).

As explained by the team member in the vignette above, in order to measure user behaviour through analytics, the team defines customized 'events' in relation to the 'actions' users can take on the platform. These events are usually composed of 'action-object' relations linked to a particular 'user interface'. For instance, 'starting a playlist on the explore page' is a customized event defined by the team. So, when a user takes this action, the related event is triggered and a *notification* is sent to developers in the form of a 'log'. In this way, it becomes possible to measure how effective a particular functionality (e.g., explore, search, etc.) is in engaging users. This illustrates how, in addition to their capacity to encode user participation, stylized social media actions are used to monitor and measure user behaviour in

order to understand how well the design choices perform in engaging users. In the following vignette, a team member illustrates these points:

> *KissMetrics works in the same way. You create events there and customize them. I mean, you define the event according to what you want to do. KissMetrics is more detailed than Google Analytics; you measure user behaviour however you want to. I'll show you. For example: for the dates that we chose here and the events that have been triggered for playlists, 87 playlists have been shared; 32 of them were started (played/listened) on the 'explore' page; some of them started in 'on boarding'; five of them were started by going to someone else's profile; two of them were started from the 'newly added playlist' section etc. We can see them all here. This way, we become able to see how the product works and we say, "Oh, even though 'suggested' is the first page, 'explore' works much better than 'suggested'. So there must be a problem with the 'suggested' page. Then we say, "Hmmm maybe we shouldn't do it that way, or we need to try something else." Then we test this new way. If it looks like it will work better, then we deploy the new feature… There's such a cycle. We use the numbers always for improvements as such. We look at what works well and what doesn't, and then try to see how we can improve a functionality that doesn't work well* **(**Erdem 11/07/2014**).**

Conversion rates are one of the most relied upon metrics for the Cubic.fm team in monitoring and measuring the performance of their design choices. As briefly defined before, 'conversions' are the set goals regarding expected user behaviours, and the term 'conversion rate' (CR) is the accomplishment level of a certain goal as defined by the team. In other words, it is the percentage of users who take a certain action that is desired or particularly expected and predefined by the team. CRs are the main reference points showing whether the system works in desired ways. In this regard, they function as one of the key performance indicators (KPIs), helping the team to optimise the platform's design by answering questions such as: 1) 'Are users are undertaking expected actions?' 2) 'What kind of optimizations should be made in order to make users undertake the required actions?' and so on.

The team produces weekly reports with the analytics tools in order to regularly monitor the overall performance of the platform. One of these reports shows the *weekly net change* of the number of users. Also known as *cohort analysis*, the report is prepared by deducting the number of 'churned' users (users moving away from the platform) from the sum of the

'newly acquired' and 're-activated' users. As aforementioned, these are the 'user states' that help the Cubic.fm team to make sense of different kinds of user behaviour based on engagement levels.

User states are used in analysis and reporting activities to gain insight into user behaviour in relation to platform features. They are often derived from the AARRR[7] framework. AARRR is an acronym delineating the five key steps for desired user states on a platform: 1) Acquisition, 2) Activation, 3) Retention, 4) Referral, and 5) Revenue. Companies customize user states according to their own key performance metrics and related expectations from their users; that is to say, criteria used to define a particular user state – i.e., being an active user, a reactivated user, and a retained or a churned user - may differ for different platforms. To illustrate, the 'churn' interval for Cubic.fm is defined as four weeks. A user who does not use the system as of the beginning of the fifth week is regarded as a 'churned' user, and her state is updated accordingly. But if the churned user starts using the system again, her state is updated as 'reactivated'.

Moreover, cohort analysis supports the measurement of user behaviour with reference to specific design features of the platform. It allows the observation and identification of which features are used and which are not used in particular stages of the user journey on the platform. In this regard, by allowing the team to monitor and measure the behaviours of a 'known sample', cohort analysis helps them to identify the characteristics of different user states in comparison to each other, thereby helping them to optimise design by augmenting features accordingly. A team member reflects on this as the following:

> *Cohort means sample. I mean, a known-sample. A sample that you know the details of, you know whom it is composed of. The purpose of cohort analysis is to monitor your (platform's) weekly growth. You need to be able grow weekly. You start to see where a known-sample you picked for this week reaches in three weeks, where the previous one has reached this week etc. … Here we look at two things: 1) A user visited the product; when is the next time she visits it again? If she visits it in the same week, what is this average time she spends here? Etc. 2) The user visits the product; then we look at whether she has tagged any playlists. These are the things we monitor*

[7] AARRR is a widely used framework by start-ups to make sense of user behaviour. See https://blog.kissmetrics.com/startup-analytics/

> *in weekly basis … How many people do I draw to the system? How many people have landed on the product? This is the traffic of our website, am I able to activate any of the users from this traffic? This activation (state of being an active user) changes from product to product. For example, on Twitter, this could be a user following 7 people, but for us it could be the user listening to music or tagging a playlist. Retention: let's say candy crush; does the user use the product and use it again within 24 hours? Because you know, you've got lives there, and the next day, you get new lives. And for us, the retention is the state where you have user tagging the playlists – just so she's engaged – and also coming back to the product to listen to music. And referral is the state where the user invites other people to Cubic.fm…Once you got the referral right, the product begins to grow on its own* **(**Baris 18/07/2014).

A negative net change in a weekly cohort analysis signals a problem regarding the efficiency of the design features in engaging users and raises concerns for the platform's growth and scale. As explained in the vignette above, even if the number of registered users is high at a particular week, unless the same users are engaged and retained on the system in the following weeks, the net change of the cohort will be negative. Thus, a weekly negative net change in cohort analysis helps the team understand how well the platform is scaling. As this illustrates, the Cubic.fm team uses analytics to gain insights about the overall performance of the platform by continuously analysing users' engagement and growth rates. Analytics, thus, provides insight by helping the team to: 1) identify the problem and 2) measure user reactions to changes made in the design.

So analytics tools play crucial roles in Conversion Rate Optimisation; that is, the optimisation of both the design of the platform and the resultant user experience. First of all, monitoring and measuring user behaviour helps the company identify problems. Tweaking the design (i.e. changing some of the features or functionalities) is the main means for addressing these problems. Analytics tools are *also* used in observing and measuring users' reactions to the changes made in platform features. If user reactions indicate that the changes fail to address the observed problem (i.e. engaging users), then another solution is architected, deployed, and tested in the same ways. One of the team members explains this as follows:

> *As a result of what we track, we understand what we need to change and how to make it work better, etc. Let's say such and such number of people out of 10 is tagging their playlists. Then we've done a particular change and now more people are tagging*

playlists. So we say, "Okay, let's keep this feature. Let's remove the other one, etc." What we're doing right now are such optimizations **(**Erdem 14/08/2014).

6.3.2. Using Analytics for Conversion Rate Optimisation

i. Using Analytics to Spot Problems and Optimise the Platform Design

Most of the observed problems - such as: insufficiency in number of playlist creators; inefficiency of the 'suggested' page compared to the 'explore' function in engaging users; or inefficiency of the 'feed structure' in creating engagement, etc. - were all identified by monitoring and measuring user behaviour through analytics. Adding the auto-tagging feature into the playlist creation process is a case in point to illustrate how analytics are used in spotting a problem in a particular activity process, and thereby how these tools are used in optimising the platform to address this problem.

It is important to be aware that the playlist creating process was not a strictly structured activity from the very beginning. Structuring this process was considered to be a feasible solution for the lack of standardized and contextualized content and data generated by users. Structuring the playlist creation activity through the process called the '*Road to the Perfect Playlist'* allowed Cubic.fm to break the process into concrete tasks to be completed by users. For example, explicating that, in order for a playlist to be eligible for sharing, it needs to have at least 10 songs in it, and at least 3 tags, ensures that each playlist includes a minimum number of songs and tags in it. Embedding game mechanics and making it a rule-based process helped to structure and standardise content and data generation.

However, structuring this process did not create the expected result. The number of playlists was still not enough to facilitate consumption efficiently. So the team focused on analysing the playlist creation process in more detail using analytics in order to understand where exactly they were losing users, since that is where the problem lies:

On the grand scale, we saw that the number of people creating playlists was significantly less than the people listening to them. The number of playlists didn't seem enough to feed the people who just come here to listen music. Like I said, we wanted people who would create playlists. Actually, we were making all these developments by looking at the data. Like I said, we brought in the automatic tagging after seeing where people were having difficulty in creating playlists… people go to

the playlist creation page and when they enter the title of the playlist, we send an event (notification) to the system. When they write the description, we send another event. When they enter the tags, we send one more event. When they complete it, we send another event. Thus, we can monitor this process step-by-step. There you see that the user enters the title, the description, but then leaves without completing the rest and she doesn't finish creating the playlist… How do analytics shows this to us? Let's say this user started creating playlist at that time, analytics sends us a playlist creation event when she press 'create playlist' button. Then the user enters the title at this time, another analytics event is sent to us for this. If she didn't do this activity, analytics doesn't send anything[8]*. For example, I go to some user's profile... Oh by the way, we do not study each user this way. We look at it statistically. But, in the detailed version, you can see, for example, that Zeynep logged in, entered the title, entered the playlist description, but the rest isn't there. There isn't any other analytics event sent to us regarding the creation of the playlist. Thus you see where the problem lies and make necessary changes. We saw that tagging was problem, so we introduced auto-tagging… By examining user behaviour as such we were able to get to know the users better. Thus they started spending more time on the platform. They started to spend time in the way we wanted them to. We can go deeper into analysing their behaviour. This way, you see, for example, where they get stuck while creating playlists* (Ozgur 18/12/2015).

ii. Using Analytics to Measure User Reactions to the Changes in Design

As aforementioned, analytics tools are not only used in spotting problems. But they are also utilised to follow up; that is, to observe and measure users' reactions to the changes made in the design to address a specific problem. This helps the team to evaluate the performance of the tweaked or newly implemented features in terms of their effectiveness in producing expected user behaviours. One of the team members explains below how they 'measure user reactions' using analytics by looking at the conversion rates:

What I mean by conversion is to define something, a goal, and then measure to what extent this goal is accomplished. It could be anything, whatever you define. For example, we have a tag button here saying, 'tag your playlists'. Using KissMetrics, we analyse how many people clicked on this button and how many of them actually

[8] i.e. The system does not send or receive any notification regarding the completion of that step.

> *tagged their playlists. We changed playlist creation structure and introduced auto-tagging. The conversion rate of this action used to be around 10% before the auto-tagging and it increased up to 35% or something. We change something like this every week, and then we look at the effects it has on users… We go, like, "This works better. Could there be something that works better than this? Of course, there could be. So how should we build that in?" We are always trying to build cycles of obtaining data, giving reaction to the data, and then optimizing the product as such* **(**Erdem 11/07/2014**)**.

The last point in the vignette above indicates that optimisation of the platform requires constant interaction between the 'platform' and its 'users'; that is, in order to optimise the platform system, the team needs users to carry on using the system and actually react to the changes made in the system, so that the team can be able to measure how users react to the changes implemented as a solution to the spotted problems. As explained by its team members, Cubic.fm relies on cycles of: 1) obtaining data on user behaviour; 2) reacting to this data by using it to spot a problem; 3) generating solutions to the spotted problem and changing the design to address it; and, finally, 4) observing users' reactions to the changes made in the system. This points to an interesting aspect of how a platform system is optimized. Cubic.fm needs *constant* interaction between the 'platform' and its 'users' for a *cyclical* process through which both platform functionalities and user behaviour are calibrated with one another. This is an important finding pointing out the potentially performative implications of the design and functioning of social media platforms.

6.4. Conclusion

In this chapter, I presented the main findings of the Cubic.fm case study. This chapter highlighted how the design, development, and evolution of Cubic.fm as a social media platform is not a linear process. Rather, it is the result of continuous optimisation of the platform design and the functionalities of the platform system in general. The main drivers of these optimisation efforts are the challenges faced after every release of the platform. Key challenges driving the optimisation of the platform have been identified as: ordering music effectively, establishing and adjusting the core interaction, addressing problems related to data integration and standardization, and increasing desired forms of user engagement. Observing and measuring user behaviour using analytics are the main means employed to address these challenges and concerns. Measuring user behaviour using analytics helped the

team to identify problems regarding user-platform interaction, to act upon those problems, and to subsequently measure how users react to the ways in which those problems were addressed. In this way, the optimisation of the platform relies on cycles of obtaining data on user behaviour in order to spot problems, then generating solutions to address them, typically through a change in platform functionalities, and, finally, following up by monitoring users' reactions to the changes made in the system.

These findings highlight how a platform-based business needs constant interaction between the 'platform' and its 'users' to be able to continuously optimise the platform design in real time. As emphasized by one of the team members, optimising platform features based on user reactions is assumed to help the Cubic.fm team to organize user participation through means that result in "*users spending more time on the platform in the way the team wanted them to*" (Ozgur 18/12/2015). This highlights the role of measuring user behaviour with analytics in organizing user-platform participation.

7. Analysis: The Platformization of Music Discovery

7.1. Introduction

In this chapter, I analyse the case using the conceptual framework and reflect on the additional findings presented in the previous chapter. As this case study indicates, the process of facilitating music discovery involves the *identification*, *selection*, *framing* and *promoting* of music based on some sort of a *knowledge on music taste* of the targeted audiences, who will be guided in discovering new content or artists. As discussed in the literature, different actors were involved in these processes previously, including but not limited to: record labels, DJs, editors, reviewers working for music magazines and so on (Smith Maguire and Matthew, 2014; Wright, 2015, Bhaskar, 2016a). For instance, the famous audience measurement firm AC Nielson has long served as one of the most prominent sources *producing knowledge* on listeners' preferences and music tastes (Napoli, 2011; Aaltonen, 2011; Wright, 2015). It did so by measuring radio listening and the sales of music records. In this way, it supported the aforementioned actors' curatorial and gatekeeping activities. Thus those traditional actors performed cultural intermediary roles and mediated the process between the production and consumption of cultural content like music. In this regard, social music discovery is chosen as a specific domain to investigate how these processes are altered by social media through the platformization of cultural intermediation.

The chapter is structured as follows: In the first section, the empirical data is analysed in light of the preliminary conceptual framework outlining the platformization of cultural intermediation. Following the conceptual framework, I first focus on the overall design logic of Cubic.fm and examine how cultural intermediation processes - namely curation, guidance and measurement - are *internalized* by the platform. I demonstrate how user-platform participation is organized around the core interaction of the Cubic.fm platform: the creation, curation, and consumption of user-generated playlists. Then I examine the three key platform functions and illustrate transformations pertaining to curation, guidance, and measurement. Secondly, I analyze the data-based operations underpinning these processes by explaining how a platform: encodes users' interaction with music, constructs digital representations of listeners in the form of aggregated data, and computes listeners' music tastes to be used in producing personalised recommendations. This part is complemented by analysis of the

findings presented in the previous chapter. Finally, the chapter concludes with further development of the ideas about how social media plat-form cultural intermediation.

7.2. Developing a Social Media Platform

7.2.1. *Design*: Core Interaction and Platform Functions

As the empirical narrative illustrates, Cubic.fm's core interaction has evolved over time in the course of developing a fully functioning social media platform. At points, it has taken a completely new form. Therefore, there is not a single, fixed core interaction that can be analysed. This paved the way for a new unit of analysis: the *process* through which the core interaction has been redesigned and optimized, along with the *reasons* behind the lack of a stable core interaction. These aspects will be examined in the second part of the chapter. Before proceeding to that part, this section first analyses the basic features of Cubic.fm's core interaction. This is done by focusing on the second version, in which the platform took the shape of a proper social media platform that was reliant on user-generated data and content.

Cubic.fm's core interaction is composed of three key activity types: the *creation*, *curation*, and *consumption* of *user-generated playlists*. The 'participants' are identified as the playlists' creators (producers) and their listeners (consumers), while the 'value units' constituting the core of user participation are the playlists themselves. In this set-up, user-platform participation is organized through a set of stylized actions (such as 'like', 'tag', add', 'follow' and so on), which constitute the basis of the platform's core interaction.

Deconstructing the activities that make up Cubic.fm's core interaction help to illustrate how user-platform participation is organised. For example, the '*Road to a Perfect Playlist'* delineates 5 basic steps to follow in 'creation' activity. These are presented in the form of 'calls-to-action': 1) 'create' playlist, 2) 'add' songs, 3) 'tag' playlist, 4) 'add' cover picture, 5) 'save as draft' or 'publish'. This set of structured actions constitutes the *creation* pillar of the platform's core interaction. As another example, user reactions are structured through a set of stylized actions (e.g., 'like' and 'up-vote'). This capacity to metricise user reactions enables them to be used as assessments of each playlist's appeal, thereby serving as a form of quality control. Thus the organisation of user-platform participation around user reactions constitutes the *curation* pillar of Cubic.fm's core interaction. Finally, the *consumption* pillar predominantly includes the ways in which users interact with the playlists and their songs (e.g., 'follow' a playlist or 'play', 'skip', and 'shuffle' songs in a playlist).

7.2.2. Incentivising Users, Facilitating Discovery, and Matchmaking

The Cubic.fm case study shows that the design logic of social media platforms is indeed built on designing a core interaction constituted of structured actions. These actions allow users to create, curate and consume content in ways that are facilitated by the platform. The analysis of the empirical narrative also reveals that Cubic.fm performs the core platform functions outlined in the conceptual framework. I will now elaborate on each of these functions and analyse how the platforms incentivise users for participation and facilitate their interactions through its matchmaking operations in detail.

i. Creating Feedback Loops to Incentivise Activity

Findings reveal that the formerly analysed platform design embeds particular features and functionalities that serve as 'incentives'. The logic described by the developers reveals that these incentives are embedded in the platform design in specific ways to keep users active and engaged by creating feedback loops. Cubic.fm incentivises user interactions using the following means: First, the platform incentivises consumption by guiding users for music discovery through *personalized recommendations*. Second, it *metricises* user *reactions*. This allows the platform to incentivise curation by rendering the assessment of content fun and engaging. This complements the ways in which creation and consumption are incentivised; the platform motivates users to assess the quality and appeal of the playlist in ways that guide the consumption of these playlists. In doing so, the platform also motivates playlist producers for further production by *notifying* them about others' *reactions*. Thus, every time a user starts 'following' a producer's playlist, 'likes' a song in one of her playlists, and/or 'adds' a song from these playlists to her own, the producer of the playlist receives a notification that informs her about how other people interacted with her curatorial work. This creates a continuous feedback loop, the logic behind which is described by one of the developers as follows:

> *The playlist-sharing version used to have more social aspects. There were people creating playlists and people listening to playlists, and the platform was fed by both of these groups of people. The person producing playlists would become content with the feedback she gets from the people who listen to, 'like' and follow her curations. Thus she would continue creating playlists and improving the quality of the content, and the*

person who listens to these playlists would come here to listen to quality music anyway. On one hand, we were trying to help one party to access quality music, and, on the other hand, helping the other party find quality feedback **(**Ozgur, 18.12.2015).

The vignette above illustrates the rationale behind the efforts put into creating self-regulating and self-reinforcing feedback loops. These were carefully designed to facilitate user-platform participation due to their supposed effectiveness in motivating users for continuous creation, curation, and consumption. This indicates how various feedback mechanisms are purposefully embedded into the platform design to increase user activity through the means of *personalisation, metricization, and notifications.*

ii. Facilitating User-Platform Participation

Cubic.fm facilitates the creation and consumption of playlists via the infrastructure it provides to its users. Using the platform system, users enact the activities related to the curation, assessment, evaluation, revision, and/or appreciation of content in ways that are facilitated by the platform through its distinct design. For example, the playlist creation process called the '*Road to a Perfect Playlist'* is a means provided to users to *identify*, *select*, *group* and *frame* songs in accordance with a particular logic (i.e. genre, mood, activity, etc.). These constituent steps were mandatory to complete and they were involved in how users curate the playlists showcasing their music taste. This illustrates the facilitating role that Cubic.fm's platform design assumes in this process: when creating a playlist, users must select and group songs in a particular order. In addition, they must sort them out by tagging the playlist and adding a visual image to further contextualize the playlist. In this way, they *curate*, *frame* and *present* music in ways that are permitted and facilitated by the platform.

The mandatory steps of this process prescribe these activities through these structured actions. Only after completing these compulsory steps, producers are able to share their playlist to help audiences (the so-called consumers) in music discovery. This indicates that the platform facilitates the creation and sharing of users' personal curations in highly controlled ways. In this way, it ensures that the desired forms of actions, which are required for the operations of the platform, are continuously performed.

To illustrate, users have to tag the playlist to make music identifiable and findable for the platform system for matchmaking operations. Likewise, users are asked to add at least 10 songs in a playlist to ensure that the playlists circulated on the platforms supposedly have rich

content. Moreover, users assess the quality and appeal of music playlists through the means built through a similar logic. Users show their appreciation of a playlist or its producer through the platform's stylized actions, such as: 'like', 'follow', 'upvote', and so on. Thus metricised, user reactions help other users to have an idea about other users' assessment of the assumed quality and appeal of the content and their producers. These metricised user reactions constitute the basis for creating aforementioned feedback loops as they also give producers an idea about the audience of their curations by providing a basic level audience measurement (i.e. the size and engagement level of their audience are provided in the form of the number of 'likes', followers, etc.).

iii. Matchmaking

Another means used to create self-regulating feedback loops are the personalised recommendations. As illustrated before, one of the main reasons to provide these suggestions is to increase the appeal of the platform by helping consumers in discovering new content that is 'matching their music taste'. This is accomplished by facilitating users' encounters with the 'relevant' content as the first thing they see on their home every time they log in to the platform. They are thus framed and presented as personalised playlists assumed to create some sort of a feedback loop. Supposedly, as they are taken up by users, the system will 'know' more about their taste and will be able to make even more relevant recommendations. This is because both the user reactions, as well as their absence, are taken as indicators of their tastes.

Constructing the personalized recommendations for listeners is accomplished by digitising their music tastes. As a result of this matchmaking process, the platform promotes particular playlists to specific users based on the computed relevance between the playlists and digitised music tastes. In the next section, I analyse the process in which music taste is rendered digital data.

7.2.3. *Data Operations*: The Process of Digitising Music Taste

In this section, I analyse the data operations underpinning Cubic.fm's matchmaking function. I do this by examining how data inputted into the backend operations are produced through the structured actions at the frontend. Thus I show how Cubic.fm encodes listening, aggregates listeners, and digitises their music taste for computing personalised recommendations. As the empirical narrative illustrates, regardless of the form that

Cubic.fm's core interaction takes, the platform architecture always shapes what users can and cannot do on the platform. For instance, the type of user participation that was facilitated by the feed structure in the crowd-sourced radio version of Cubic.fm was substantially different from that of its subsequent playlist-sharing version. Whilst in the first version users could only share *a song*, the second version only allowed the sharing of *playlists*.

Moreover, criteria were established so that users were only allowed to do so by completing all of the necessary steps required to publish (share) a playlist. Furthermore, because the latest personal music library version was designed specifically for personal use as a stand-alone product, it was not even possible to interact with other users in that last version of the Cubic.fm platform. These points illustrate that the user-platform participation on Cubic.fm is strictly conditioned by actions constituting its core interaction. This shows how the platform design gives direction to user behaviours, rather than just registering them as they naturally occur. This provides evidence for how social media engineer user-platform participation, thereby engineering the resultant user experience. In the next section, I elaborate on the logic behind the ways in which the Cubic.fm platform prescribes user behaviour to measure their music taste.

i. Encoding Listening

The conceptual framework outlines the logic behind the ways in which user-platform participation is structured by the core interaction of a social media platform. This is also shown to be the case for Cubic.fm. As a social media platform, Cubic.fm strictly prescribes the ways in which they can interact with music on the platform. This is accomplished through a set of stylized actions (i.e. 'like', 'add' 'create', 'tag', 'share', etc.), which are comprised in the core interaction of the platform. In this way, Cubic.fm encodes users' interaction with the music objects (both the playlists and the individual songs in the playlists) through such discrete, yet highly standardized user actions. Due to the supposed explanatory power of these actions as listeners' dispositions, Cubic.fm *encodes them to disassemble* listeners' music tastes and to *reassemble* it in the form of digital data. Thus it becomes possible to use it to construct recommendations personalised for users' taste.

The logic of encoding user-platform participation on Cubic.fm involves producing data about listeners by creating 'action-object' relations. As described in the empirical narrative, the main 'objects' of the platform system are the following: *Users*, *Playlists*, *Artists*, and *Tracks*,

as well as the Tags attached to these tracks. Aforementioned structured actions, constituted of anything a user *can* do on the platform, link these 'objects' to each other. Activities pertaining to the creation, curation, and consumption of music - such as 'creating-a-playlist', 'adding-a-song-to-a-playlist', 'tagging-a-playlist', 'playing-a-playlist', and 'liking-a-song-in-a-playlist' - all illustrate how Cubic.fm produces data about listeners by *encoding* user-platform participation in the form of action-object relations. Disassembled as such, listeners' predispositions become eligible for the next step: constructing digital versions of listeners.

ii. Constructing Digital Listeners

Cubic.fm *disassembles* listeners' identities through encoding. Encoding users' interactions with music in the form of the aforementioned 'action-object relations' allows the platform decompose listeners' predispositions and turn them into discrete data tokens. These data tokens are then normalized as 'user-tag' relations, thereby becoming even more granular. Finally, these tags are piled up to *reassemble* listeners by constructing the digital version of their identities in the database. This is what Cubic.fm developers refer to as constructing 'Music IDs' of listeners (See Interview #20 for an illustration.). The platform re-constructs listeners in its database in the form of 'digital data objects' and thus renders them amenable to computation. As a result, digital listeners are produced.

Normalization plays a crucial role in this process. As explained in the empirical narrative, encoded data enters into Cubic.fm's database after being normalized. This is accomplished using the tags of the object comprised in an action-object relation. In this way, action-object relations are decomposed and become further standardized to take the form of *a* user's relation with tags of these objects.

This shows that the relational structure of Cubic.fm's database is built on a particular logic that sees 'each action of a user' in connection to the 'tags' of the 'objects' of these 'actions'. To illustrate this, we can draw on the previous example given while presenting the case. Let's say a user, *Mary* likes a song from Bob Marley. Encoding this activity in the form of an 'action-object' relation includes the 'action' as 'like' and the 'object' as the song called 'NoWomanNoCry'. This object involves the following tags: Reggae, Jamaica, Rastafarian, and the 70s. So, when Mary likes this song, her action is encoded by Cubic.fm and takes the form of the following action-object relation: Mary-likes-NoWomanNoCry. In the second step, encoded data is normalized using the *tags* of the *object* of this *action*. Thus the resultant

data take the form of the following 'user-tag' relations: Mary-Reggae; Mary-Jamaica; Mary-Rastafarian; Mary-70s. In line with the logic of the relational database structure, these 'relations' are represented in the database by creating a link between *User* and *Tag* objects in a relational table called the 'User-Tag-Score'. This process is *repeated* for each tag of an object involved in an action-object relation, and this table is updated accordingly every time a user takes an action on the platform. Consequently, digital music identities of users, the so-called 'Music IDs', are constructed in the form of aggregated tags.

iii. Digitising and Computing Music Taste

The aggregation process is followed by the computation of taste. This first entails insertion of the normalized data into the Cubic.fm's database in the following way: For each tag involved in an encoded action data, a new 'user-tag' object is created in the aforementioned 'User-Tag-Score' table. This connects *User* and *Tag* objects. Each user-tag object inserted into this table (e.g., Mary-Reggae, Mary-Jamaica, etc.) is given a value based on the predetermined *weight* of the encoded *action.* This varies depending on its *type* and *recency* of the actions (i.e. the weight of *tagging* action is 8, whereas the weight of *listening* to a playlist is 3). If there is already an existing object in this table created for a particular tag before, then its value is updated in the same way. Thus the tags of the encoded actions are aggregated, quantified, and scored in this 'User-Tag Score' table. Calculated as such, respective tag-scores are then ranked, and the top ranking tags are accepted as signifiers of a listener's music taste. Through this process of digitising and computing music taste, Cubic.fm constructs personalized recommendations.

This process also involves *identifying* the relevant music *content* matching users' tastes. Cubic.fm profiles playlists using the similar means involved in digitising music taste. Findings show how Cubic.fm, as a social media platform, outsources the task of 'sorting out' playlists to its users. However, in its early crowd-sourced version, this task was performed by the platform itself, using methods like signal processing. Adopting the social media model made Cubic.fm reliant on its users, not only for user-generated data and content, but also for the now-outsourced task of categorizing music. In this way, it became the task of 'producers' on Cubic.fm to make sense of the music content. They are the ones who select, group, and order the songs. They must also contextualize the playlist by adding labels that describe the commonalities shared by the songs grouped together in a particular playlist o (i.e. mood, genre, origin etc.). This is accomplished by 'tagging' the playlist.

Users are also required to add a cover picture for the playlist to give visual clues to the 'consumers' (listeners) about the playlist content. However, in contrast to contextualising playlists with a cover picture, 'tagging' not only guides listeners, but it also guides the platform system in identifying the *content* to be *matched* with *listeners*. It is through these means that the Cubic.fm system processes music playlists as *value units* to be used in matchmaking. In addition to the user-generated tags, Cubic.fm also uses metadata aggregated from different sources, such as Last.fm and the EchoNest. These data help the Cubic.fm system to calculate what can be referred to as 'item-tag-scores'. I have discussed how the system computes music tastes by ranking the calculated scores of user-tag relations. Playlist tags and the metadata aggregated about the songs in a playlist are scored in a similar way. Then these are ranked to construct a tag-profile for playlists. This enables Cubic.fm to present personalized suggestions to users in a *particular order* based on the computed strength of the 'relevance' between the playlists and a users' music taste.

The system computes the 'similarity' among 'tag-profiles' constructed for *users* and *playlists* by calculating the 'distance' between them. Playlists that have high scores for the tags, which are accepted as the signifier of a user's music taste, are placed on the top of the personalized list in which suggested playlists are shown to this user (i.e. playlists that have the higher scores for the top-ranking 5 tags in the user-tag-score table). These aspects reveal how data operations underpinning the platform's matchmaking operations are built on a set of assumptions about what taste is and how it can be measured.

Consequently, as this analysis demonstrates, these research findings validate the preliminary framework outlining the platformization. However, findings also reveal that the design of the core interaction of a social media platform is not a straightforward process. This case study showed that the core aspect of how the platform organizes the user-platform participation is determined by a basic design rationale. Then it is continuously fine-tuned to optimise the ways in which the platform facilitates user interactions. The next section elaborates on this optimisation process to refine and further develop the ideas put forward when devising the initial conceptual framework.

7.3. Optimising a Social Media Platform

I analysed the relationship between the frontend and backend of the Cubic.fm platform based on 2 key dimensions: *platform design* and *data operations*. As this analysis shows, the

findings of the Cubic.fm case study correspond to and further validate this proposed framework. However, findings presented in the previous chapter reveal that there are additional aspects that should be taken into account when explaining the platformization of cultural intermediation by social media platforms. These are especially related to the issues and concerns regarding the ways in which a social media platform is optimized after its initial design and development. The Cubic.fm case study has provided a valuable opportunity to examine these aspects. In this section, I take this opportunity and draw upon its findings to reflect on the aspects that are not covered in the preliminary conceptual framework. In so doing, I further develop my ideas to form the foundation for the Discussion chapter.

A close look at the findings presented in Chapter 8 reveals that the development of a social media platform involves several challenges. These include, but are not limited to: ordering value units used in matchmaking operations, building a standardized data infrastructure, adjusting the core interaction, and increasing user engagement to secure data and content production. I demonstrated how these challenges play significant roles in organising user-platform participation and how it is adjusted over time.

I then explained how analytics tools and accompanying measurement techniques are used to monitor and measure user behaviour on the platform to address the aforementioned challenges. Their usage is shown to be twofold: to spot a problem and to observe how users react to the changes made to address that problem. I first examined how analytics tools are used to spot problems regarding expected/desired user behaviour (conversion). Then I discussed how these spotted problems are addressed, and subsequent user reactions are measured to find out whether the deployed changes effectively created the desired behaviour (conversion optimisation). It is thus shown how, using analytics tools in this way, the platform design is continuously attuned to adjust the core interaction.

In this light, the Cubic.fm case study shows that the development of a social media platform is not a linear process in which the development precedes the launch of a fixed or a stable system. On the contrary, the findings provide evidence to conclude that it is an ongoing dynamic process constituted of cyclical data operations through which the platform design is continuously fine-tuned. This confirms previously mentioned views that suggest conceptualising social media platforms as 'dynamic objects' that are "*tweaked in response to their users' needs and their owners' objectives, but also in reaction to competing platforms and the larger technological and economic infrastructure through which they develop* (van

Dijck, 2013: p.7)". This latter point is also evident in the Cubic.fm case study, and elaborated on in the next section.

7.3.1. Building a Unified Data Infrastructure

Findings reveal that social media platforms are 'tweaked', not only in accordance with and in reaction to users and owners, but also in relation to the requirements of other services, and, more generally, *the ecosystem within which they operate*. Challenges faced by Cubic.fm while building a standardized data infrastructure illustrate this well, namely: the dependency on too many APIs and the contingency arising from being dependent on other services.

According to the findings of this study, a social media platform like Cubic.fm significantly relies on external parties (i.e. streaming services like Spotify, on-demand services such as YouTube and SoundCloud, metadata services such as Last.fm or the EchoNest, weather forecast services as used in the first version, and so on.). For this reason, it is subjected to their requirements and constraints for data aggregation and formatting. These aspects indicates the importance of building a continuously developing 'data infrastructure'. As the empirical evidence indicates, building this database relies on the encoding mechanism underpinning data production on the platform and the data aggregated from other sources, which undergo a standardisation and unification process through several rounds of *normalisation*. This data aggregation process may also require feeding platform-generated data to the services as part of the rules and the requirements specified in their APIs for data exchange and/or aggregation.

Data synchronization between Cubic.fm and streaming services illustrates this point well. Previously, I described how the social discovery version of the platform was built on a structure that allows users to import pre-existing playlists that they created on other music platforms. That model allowed the system to automatically pull in the playlists previously created on the other streaming services when a user connects her streaming service account to Cubic.fm. Playlists were pulled in together with the behavioural data of the user that was aggregated by that other service. From that moment onwards, whatever action a user takes on Cubic.fm syncs back to the streaming service and vice-versa. This model, dependent on *data synchronization* between these platforms in ways specified in their APIs, indicates that APIs are highly crucial elements of data infrastructures underpinning the core operations of social

media. Building this infrastructure is observed to be an ongoing process; it constitutes a significant part of platform optimisation, and, thereby, platformization.

7.3.2. The Cyclical Relation between Data and Analytics

Chapter 8 provided a detailed account of the challenges and concerns involved in optimising the platform as well as the ways in which these are addressed. To briefly summarise, using analytics tools requires a continuous stream of user data input to monitor and measure user behaviour. In turn, user-platform participation and, consequently, *the ways in which behavioural data is produced* are continuously fine-tuned to *shape* the resultant *data* in *desired ways*.

The underlying reason for this is to optimise the platform system and satisfy the user needs and business requirements in ways that pave the way for scaling the platform through network effects. It was discussed how this creates a cyclical relationship between *data* and the use of the *analytics* tools that are fuelled by the data. This continuous optimisation of the platform design indicates that data analytics also constitutes an important aspect of platformization. Data analytics tools allow developers to experiment with new functionalities, while enabling them to monitor, measure, and analyze users' reactions towards newly implemented features. It thus enables them to modify these features accordingly if/when necessary. Furthermore, the data about user behaviour are generated through preformatted features as part of the platform design, and then they feed back to the optimisation of the very design itself. This mechanism further shapes user-platform participation and the resultant data generation in cyclical ways, thereby paving the way for a *mutual calibration* of the *platform* and the user *behaviour*.

7.4. Explanatory Framework

This analysis of the platformization of music discovery provided an empirical basis to refine the preliminary ideas put forward to explain platformization of cultural intermediation. The analysis reveals that platformization is significantly linked with how a social media platform is *developed* and *optimised*: 1) The development of a social media platform involves two core aspects: preliminary configuration built on a particular design logic and the data operations underpinning the platform's matchmaking function; 2) Optimisation builds on this initial configuration and results in mutual calibration of the platform and user participation due to a

continuously developed data infrastructure and the cyclical use of data analytics. A crucial point to recognise is that these elements of platformization are closely intertwined and dynamically reliant on each other. Their relationship is summarised in the next section.

7.4.1. Development

The design of a social media platform follows a particular logic: bringing together external producers and consumers of content and facilitating their interactions around the design of a core interaction (Parker et al., 2016). This core interaction is constituted of several actions that can be broadly subsumed under three key activity types: the creation, curation, and consumption of user-generated value units (i.e. cultural content like music playlists). Cultural intermediation processes – namely, *curation, guidance* and *measurement* - are internalised by the platform in the form of the aforementioned activities that make up the platform's core interaction. This set-up enables the platform to perform its core functions while also being sustained by them.

The platform organizes user-platform participation and facilitates interactions using pre-formatted *stylised actions*, which are recognised in the literature as unique aspects of social media platforms' design (Gerlitz & Helmond, 2013; Alaimo and Kallinikos, 2016; 2017; Van Dijck, 2013; Couldry and Kallinikos, 2017). They also allow the platform to facilitate interactions by incentivizing and controlling platform participation through the strategies of *personalization*, *metricisation*, and *notification*. These are embedded in platform design in order to create *self-regulating feedback loops*. The same design features also enable the encoding of user-platform participation in the form of standardised data tokens, which allows commensuration of different actions (Alaimo and Kallinikos, 2016; 2017). Generated in this way, discrete data tokens representing user behaviours are then aggregated to construct digital versions of the identities of individuals *qua* users. This aggregated data is then used to compute their 'digitised taste'. Thus the so-called 'user taste profiles' are produced to be used in the matchmaking operations of social media.

This distinct set up, which is built on the entwinement of design and data operations, allows a social media platform to *internalise* curation, guidance, and measurement in ways *shared out* by the platform and its users. The platform provides the content producers the means through which they selectively bring together and curate the value units (i.e. playlists) that get consumed on the platform. These value units are produced through the curatorial work of

producer-users (i.e. identify, select, frame, and promote music tracks to curate a playlist, etc.) in ways structured by the platform design. These value units are then made available to consumers through the platform's matchmaking operations. Personalised recommendations are the key aspect of this matchmaking process, in which this time the platform itself performs curation: it *identifies*, *selects*, and *curates* a list of user-generated value units (i.e. playlists) to be presented to the consumer as the value units matching their tastes. This indicates that consumers are guided to discover new content in novel ways thanks to: i) the content producers' showcasing of their tastes through their personal curations and ii) the curatorial work performed by the platform system to *match* those personal curations with the 'right' audiences. As a result of this matchmaking, consumers/audiences are selectively presented with a set of user-curations, and they are hence *pushed* towards particular content but not others.

This form of guidance via a combination of platform and user-led curation is complemented by the quantified user-reactions. Enabled by the structured design features, these quantified reactions function as a novel assessment mechanism that guides consumers about the quality, appeal, and popularity of the user-curated content. At the same time, they function as a primitive form of audience measurement and are strategically used to incentivise production as part of the aforementioned feedback loops. Simultaneously, the platform performs another form of audience measurement underneath the surface in the course of digitising and computing taste to be used in matchmaking operations.

7.4.2. Optimisation

The process of platformization is not completed with this *development* process. It also entails the continuous optimisation of a social media platform following the initial set-up of the platform's core. The optimisation process entails constructing and continuously improving a data infrastructure while constantly monitoring the platform activity using data analytics to fine-tune the platform design. In combination with the encoding mechanism used in data generation, APIs play crucial roles in structuring data and the formation of a standardised, unified database upon which the platform performs its core functions. This renders APIs a central component in the process in which data is rendered platform-ready (Helmond, 2015), enabling data generated on one social media platform to be disseminated across the web and variously used by other platforms. APIs are comprised of specifications regarding the standards, data formats, rules, and other requirements for data aggregation and

synchronisation among different platforms. Hence, just like the encoding mechanism, using APIs for data aggregation to profile users and content can be considered as a significant aspect of platformization.

This optimisation process also relies on using data analytics tools, which requires constant interaction between the *platform* and its *users.* This is the main criterion for using the data analytics effectively to optimise and fine-tune the platform system in ways that increase user participation. In order to fuel analytics tools, a platform needs an influx of data generated through user activity (Couldry and Kallinikos, 2017). This data influx can only be generated as long as users stay *active* on the platform. To enable this, social media platforms strategically design and optimise feedback loops that afford facilitating user *actions* and *reactions* in self-regulating and self-reinforcing ways (Pellika, 2014; Gerlits and Lury, 2014). User-platform participation is organised around these very same structured actions and reactions, through which the platform encodes user behaviour to digitise their taste and perform matchmaking operations.

Data analytics is a significant element of creating these feedback loops using the structured call-to-actions. Using data analytics allows the platform to: i) spot the problems in the design (i.e. problems that decreases user engagement), ii) deploy new features to address these problems, and iii) measure how users react to these new features (i.e. whether their engagement level increases or not, and, if not, going back to repeat the first process). This picture suggests that the continuous use of data analytics operates *recursively*. The design choices, which shape user behaviour, embody particular assumptions related to developers' expectations from users to behave in particular ways. If the expected behaviours are not performed, a new set of hypotheses is developed about user behaviours. Related assumptions are embedded in the new design choices and then, again, developers-analysts measure how users react to those changes. Thus the assumptions on user behaviour are continuously validated or falsified through hypotheses-testing in real time, based on the '*expected'* versus '*resultant'* user behaviours. Consequently, both the platform and the user behaviour are shaped and calibrated with each other.

The following figure visually illustrates the main aspects of the final explanatory framework.

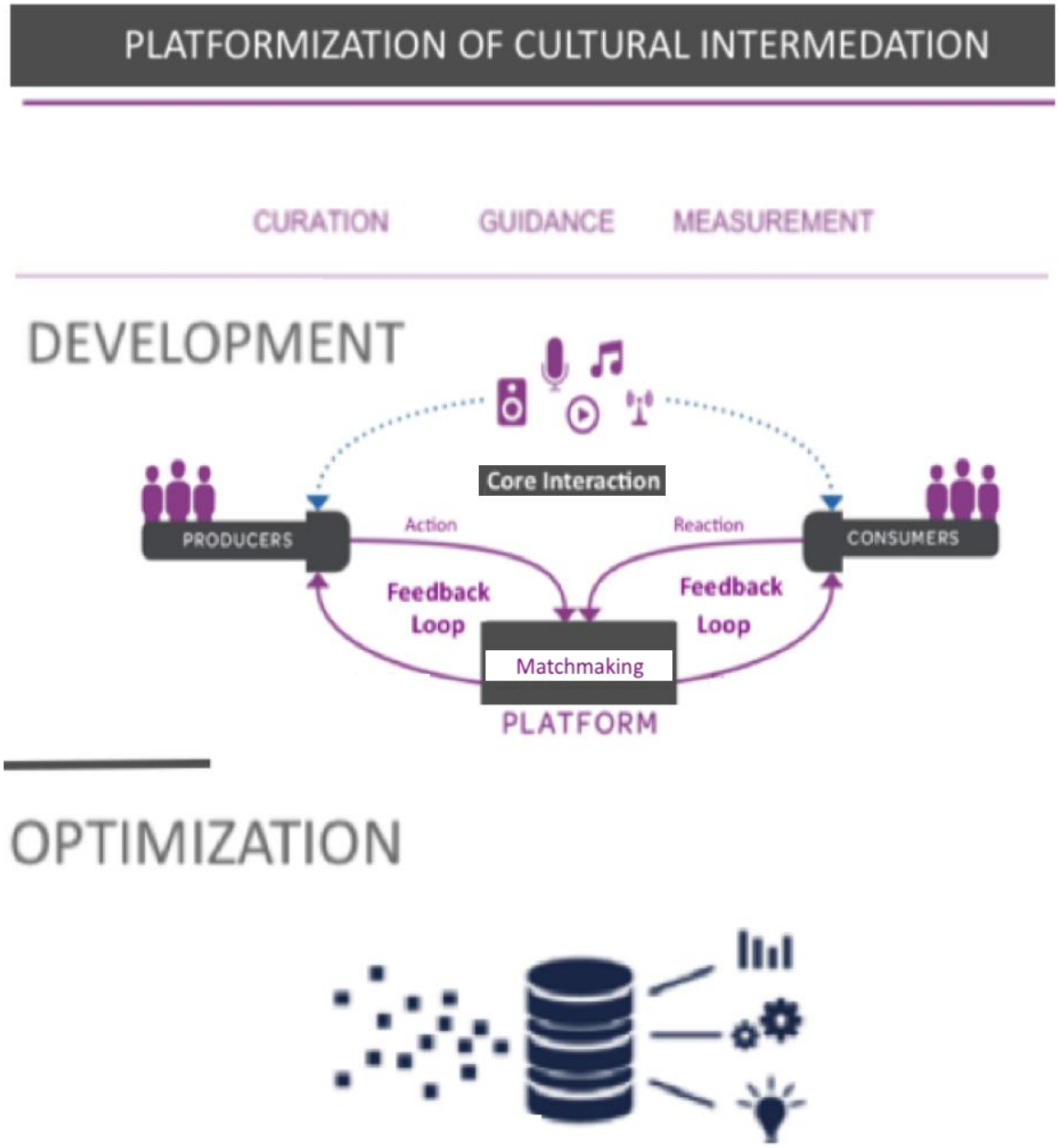


Figure 7.1: A visual illustration of the explanatory framework delineating the Platformization of Cultural Intermediation[9]

[9] The visual model is sketched by modifying the model provided by Parket et al.,(2016). Source: 2016, Parker & Van Alstyne, with Choudary - licensed under creative commons attribution-ShareAlike 4.0 International (CC BY-SA4.0). Available at: https://inform.tmforum.org/internet-of-everything/2017/04/pipes-platforms-leveraging-network-effect/

7.4.3. Platformization of Cultural Intermediation

i. Curation

A social media platform facilitates what can be considered as ‘user-led’ curation in ways that are allowed and constrained by the platform design. It structures how producer-users curate content to showcase their taste to consumers by *selecting, categorising, framing and sharing* their curations. In this setup, the relevance between the consumers and the shared-content is established by the platform. Thus particular consumers are rendered target audiences of the content shared by specific producers.

This indicates that concurrently the platform system itself also plays a similar curatorial role at another level as the matchmaker enabling the aforementioned interactions between the producers and consumers. During the matchmaking process, the platform undertakes all of the aforementioned ‘user-led’ actions: it *identifies*, *sorts out*, *selects*, *clusters*, and *presents* playlists to a specific user in a particular order and sequence. This process can be referred to as ‘platform-led ‘curation. This is enabled by digitising tastes through a set of data operations involved in constructing personalised recommendations. This process also functions as a novel kind of audiencemaking as the platform renders consumers the target audiences of the personalised content. This layered configuration, combining user-led and platform-led curation, illustrates what can be considered as ‘Plat-formed Curation'.

These findings validate previous views that highlighted how curation practices are altered by a set of new actors such as algorithms, users and social media platforms (Ashton and Couzins, 2015; Barna, 2017; Bhaskar, 2016; Bruns, 2011; Lange, 2016; Potts et al., 2008; Jansson and Hracs, 2017; 2018; Stanoevska-Slabeva et al., 2012; Villi, 2012).

ii. Measurement

The same analysis can be performed for the platformization of measurement. Metricised user actions (i.e. ‘like’, follow, etc.) constitute the basis for creating feedback loops between producers and consumers of user-curated content. Quantified as such, consumers' reactions to the content provide producers a form, albeit basic, of ‘audience measurement’. Thus informed about the size of their audiences, producers are assumed to be motivated for further content production. Thus a self-reinforcing feedback mechanism is enabled by the metricising capacity of the stylized actions. Thanks to the encoding capacity of the same structured actions, the platform itself performs more advanced forms of audience measurement

underneath this presentation layer. This process involves creating knowledge about the cultural dispositions of individuals in novel ways. This entails digitising their taste by encoding and aggregating their online behaviours in real time. This new form of knowledge production process about individuals' tastes constitutes the foundation for the platform's matchmaking operations and gives rise to a unique form of audiencemaking.

iii. Guidance

Plat-formed curation and measurement establish the basis through which social media plat-form guidance. Producer-users provide guidance for the consumers by their curations. This illustrates 'user-led' guidance. The platform presents these user-led curations to their potential target audiences based on the computed relevance of the content and the computed taste of users. Thus a 'platform-led' form of guidance is provided along with the previously outlined platform-led curation. This is complemented by another form of guidance that is provided by users through the ways in which they assess the appeal of the user-curated content with the help of the metricised actions. These stylised actions enable the audience of the content to give feedback to their producers in standardised and scripted ways while guiding other users about their appreciation of user-led curations. Thus user-led forms of guidance are complemented with the unique role that the platform performs in this process.

7.4.4. Plat-formed: Merged, Layered, and Shared

This section analysed how social media plat-form the curation, guidance, and measurement. It is important to recognize the layered form they take in this processes. Both the users and the platform perform these practices simultaneously at different layers: it is the users who perform these activities at the presentation layer, only in ways that are permitted and constrained by the platform's structured interface. In the meantime, the platform system itself performs similar activities underneath this surface in ways that are not visible to users. Thus these processes become layered in the course of being internalized by the platform and shared out between the platform and its users.

This confirms previous views on how social media's distinct logic differs from the mass media logic. Bhaskar's (2016a) explains this shift's implications for cultural intermediation by describing how the top-down industrialised organization of cultural content, the so-called 'broadcast model', is being replaced by a 'consumer-curated model'. In the latter, consumer-curated content, such as user-generated playlists, is considered to be replacing their broadcast

versions, such as albums. This indicates that tastemaking practices, such as guiding listeners on what to listen to, have shifted from experts in the media and culture industries to ordinary audiences.

Moreover, it is discussed how previous forms of assessing cultural content is significantly transformed. The quality assessment used to be provided partially by gatekeepers and partially by critics and reviewers (Wright, 2015) through the more traditional forms of guidance. This study reveals how these practices are being replaced by user curations, quantified user reactions, and personalised recommendations. In this regard, participating in curation and guidance practices renders users 'social curators' (Villi, 2012) or 'gatewatchers' (Stanoevska-Slabeva et al., 2012), performing a new form of cultural intermediary roles in ways that are facilitated by the platform. Likewise, prevalent forms of audience measurement (Napoli, 2011) and audiencemaking (Ettema and Whitney, 1994) have now taken novel forms thanks to the unique ways in which social media provide primitive forms of measuring audience size (Baym, 2013) and produce knowledge about taste (Alaimo and Kallinikos, 2017). Consequently, in the course of being internalized by social media, processes that were once in the control of cultural experts and media companies (Wright, 2015), such as measuring the size and engagement of an audience and producing knowledge about taste, are substituted by the users and the platform.

These aspects highlight how social media afford a distinct form of audiencemaking. Data representations of users' taste are crafted through encoding their behaviour and used to fuel social media's matchmaking operations. This entails rendering users target audiences for particular content based on its computed relevance to users' tastes. This is a significant shift from the non-digital audiencemaking processes. In the past, audience measurement was guided by predefined consumer segments and categories (Aaltonen, 2011; Aaltonen and Tempini, 2015; Wright, 2015), and audience behaviour was measured *during* the consumption via traditional audience measurement processes[10]. Nevertheless, audience-*making* was a *subsequent* process to this measurement of audience behaviour (Aaltonen and Tempini, 2015). Accordingly, in the past, results of the measurement and analysis of

[10] An example of this is the employment of separate rating devices such as 'People Meters' introduced by AC Nielson - a global marketing research firm best known for its audience measurement system that measures television, radio and newspaper audiences. In 1986, Nielsen developed People Meter, an electronic method that moved audience measurement from an active and diary-based form to a passive-monitoring version as it started to simultaneously record real-time viewing (Aaltonen, 2011).

audience behaviour used to *feed back* to the audiences *later*, in the form of *post-hoc*-made programming decisions (i.e. the schedule and content of the TV or radio programs). Such decisions used to shape cultural encounters by determining what audiences see on TV or hear on the Radio and when (Aaltonen, 2011).

In the course of being internalized by social media platforms, the measurement of audience taste becomes a *concurrent* process, which simultaneously takes place with the consumption of cultural content. Thus, the consumption of content and measurement of audience taste merge during social media's audiencemaking operations. However, this is not the novelty that social media brought to audience measurement, as, from the 1980s onwards, audience measurement started to be conducted while audiences were watching TV or listening to the radio. Therefore, the novelty social media brought to this process is *not* the *concurrency* of measurement and consumption. Rather, it is the generation of data about audience tastes in ways that instantly *fed back* into users' consumption decisions by shaping what they see on the screen: every time a user refreshes her home page, a new set of recommendations are shown that are based on the dynamic taste profile that the platform constructs for each user.

This analysis indicates that social media significantly transform curation, guidance, and measurement. Thus, by altering these traditional processes mediating between the production and consumption of cultural content social media operate as unique kind of cultural intermediaries.

7.5. Conclusion

This analysis utilised the concept of *platformization* as a theoretical framework to unpack social media's sociotechnical configuration in cultural intermediary terms. Empirical findings, collected from a social media platform focused on music discovery, helped to further develop the ideas that were put forward while devising the preliminary conceptual framework. The initial ideas regarding platformization are refined to develop an explanatory framework. The data-operations of the studied platform involve all the key instances of the cultural intermediation process investigated in this research. Thus, the empirical case study helped to illuminate how cultural intermediation is 'plat-formed' by social media through the internalising of curation, guidance, and measurement during the *development* and *optimisation* of a platform. The empirical case revealed how, in the course of platformization, these processes merge, get layered and shared out by the platform and its users. In the next

chapter, I discuss the implications of this research by linking the findings back to the literature that provided the theoretical basis for this study.

8. Discussion

8.1. Introduction

In the previous chapter, the preliminary conceptual framework devised for this study is applied to the analysis of the empirical data. This allowed assessing and improving the initial conceptual ideas. Thus an explanatory framework is developed to better account for how social media transforms cultural intermediation. My analysis contributes to the study of how a social media platform internalises the prominent cultural intermediation processes, namely: curation, guidance, and measurement. In order to achieve this, the notion of *platformization of cultural intermediation* is developed as the theoretical approach to unpack how social media platforms operate as matchmakers (Evans and Schmalensee, 2016) by internalising traditional value flows (Parker et al., 2016).

The concept of platformization helped explaining how a particular sociotechnical configuration gave rise to a distinct type of *intermediary logic* (Langley and Leyshon 2017; Plantin, 2018; Parker et al., 2016; Nieborg and Poell, 2018; Van Dijck et al., 2018). This intermediary logic is recognised as a significant change in how social, cultural and economic activities are mediated across industries (Parker et al., 2016), including the culture industry (Nieborg and Poell, 2018). This research illustrated this by a case study on how consumption of music is now facilitated by online platforms. This is the empirical contribution of this thesis. It is an important contribution as it helped demonstrating how the core function, and indeed the raison d'être, of these platforms is to bring together the producers and consumers of cultural content and consummate matches between them (Parker et al., 2016; Nieborg and Poell, 2018; Van Dijck et al., 2018). These matchmaking operations remain significantly understudied. This is due to the difficulty of having access to relevant empirical sites, as the proprietary algorithms used in these processes are one of the most important competitive advantages of online platforms. Because, the industry as a mechanism for generating distinction and value within the marketplace (Webster et al., 2016).

In this regard, the insights provided by this empirical study make significant contributions to the literature. The thesis also advances the existing scholarship by highlighting and empirically demonstrating how, as emphasised by the informants in the studied case, data aggregation is more important for social media platforms than their proprietary algorithms. This is proven by deconstructing the complex data work preceding the work of algorithms.

It is shown that representing users in the form of data objects is the key mechanism underpinning how personal curations of listeners are made known to other users by constructing personalised recommendations matching these listeners' tastes. Thus, it is shown how user-led curations showcasing users' taste are delivered to target audiences through the platform-led curations. This is the core of social media's matchmaking operations. Deconstructing matchmaking operations of the chosen platform illustrated this. Thus it helped explaining how the processes of curation, guidance and measurement merge, get layered and share out by these platforms and their users in the course of being internalised by social media.

In order to unpack this structure, initial conceptual ideas focused on exposing the relationship between the *frontend* and *backend* of social media. The aim was to deconstruct the design logic of the platform-based business models and the data-based operations that are unique to social media's matchmaking function. Applying these ideas to the analysis of the chosen platform allowed to explain in detail how a range of sociotechnical elements take part in platformization of cultural intermediation. This empirical case study revealed that both the *users* and the *platform-system* play crucial roles in transformation of cultural intermediation in the course of the *development* and *optimization* of a social media platform.

This chapter discusses these findings in ways highlighting how this research contributes to scholarship. The chapter is structured as follows: First, I conceptualise social media platforms as contemporary sociotechnical cultural intermediaries. Following this, I move on to discuss previously problematized overemphasis on the users and the algorithms. In doing so, I conceptualise those as part of the distinct configuration of social media enabling them to operate as cultural intermediaries. This is followed by a discussion of the implications of how users are represented through engineering devices and how this is linked to the mutual calibration of the platform and user behaviours. Finally, these issues are linked to the wider debates concerned with the implications of categorisation and measurement through a discussion on social media's role in tastemaking.

8.2. Social Media: Contemporary Cultural Intermediaries?

The role of personalised recommendations in transforming cultural intermediation is recognised by previous studies. Scholars have pointed out how algorithms, users and social media involve in curation and alter the traditional ways in which cultural content is framed

and represented (Ashton and Couzins, 2015; Barna, 2017; Bhaskar, 2016; Bruns, 2011; Jansson and Hracs, 2017; 2018). Complementarily, others highlighted how contemporary sociotechnical arrangements exercise power on shaping encounters with content in ways implicated in shaping cultural tastes, in highly subtle ways (Beer 2009, 2013; Morris 2015a; Webster et al., 2016; Wright, 2015). Some scholars consider this to be an indication of the emergence of a new infrastructure for tastemaking (Beer, 2013; Wright, 2015), pointing out a potential shift in the ontology of taste formation (Beer, 2013). In this regard, personalised recommendations offered by online cultural services are considered to define how culture is ordered, structured, sorted, and thus shaped in particular ways (Beer, 2013; Van Dijck, 2013; Gillespie, 2010).

This study conceptualises social media as *sociotechnical cultural intermediaries*, operating not only as *matchmakers* but also contemporary *tastemakers*. This is the most important contribution of this thesis. Previous studies focused on recommender systems as cultural intermediaries shaping cultural tastes (Karakayali et al., 2017; Morris 2015a;Webster et al., 2016). Thus they are considered to be a form of contemporary tastemakers (Santini, 2018). These studies provided valuable insights regarding how this new form of intermediation and tastemaking depends on the work of users and algorithms of recommender systems. This thesis builds on and extends these insights by conceptualising users and algorithms as part of larger sociotechnical structures like social media as platform-businesses. Their functioning relies on such recommendation systems, but equally importantly, it is shaped by their business requirements underpinning their distinct design logic. In this regard, this thesis makes an important contribution to the existing knowledge by introducing the concept of *platformization* as a theoretical approach to study social media platforms in cultural intermediary terms. This is a substantial contribution as it advances knowledge in this field by empirically demonstrating why future studies should take into account the broader context that created this new intermediary logic.

Social media in this study are conceptualised as prominent instances of a larger transformation conceptualised as 'platformization' (Constantinides, Henfridsson & Parker, 2018; Helmond 2015; Nieborg and Poell 2018; Parker et al., 2016; Van Dijck et al., 2018). This theoretical approach revealed that the work of users and algorithms are indeed crucial components of transformation of cultural intermediation. However, empirical evidence indicate that a better understanding of how they relate to each other can be better

built only by analysing them within the broader context in which they involve in the presentation and representation of culture (Bourdieu, 1984; Smith Maguire and Matthew, 2014).

The reason for this is twofold: First, the production of personalised recommendations does not rely solely on the work of algorithms. Prior to this, users taste should be turned into data and readied for algorithmic processing. Secondly, producing knowledge about users' taste is a highly taken-for-granted aspect of social media. It is recognised that knowledge production is significantly linked with the issues of power (Foucault, 2007). Nevertheless, despite being highly trusted sources for acquiring information about people, social media's knowledge production mechanisms are rarely questioned and investigated. Throughout this thesis, it is emphasised that social media data is *not* a neutral by-product of user activities as often assumed. On the contrary it is a carefully crafted product of social media's prescriptive design structuring the work allocated to their users. This prescriptive design shapes users' activities, which is taken as the basis to digitise their taste for algorithmic matchmaking. These data representations of taste are also shaped by the cyclical use of analytics in the course of the development and optimisation of a social media platform. This is because, the structured user activities creating data for constructing taste profiles are frequently tweaked by the platforms' developers in the course of the experiments they conduct to increase user engagement. Empirical illustration of this non-neutrality of social media data is a crucial contribution advancing existing critical views on social media's knowledge production mechanisms.

This empirical contribution illustrates how sociotechnical systems do not just depend on the work of technology or people but rather they are made-up of a complex interplay between the *people* performing specific *tasks* using *technology* in particular *structures* (Leavitt, 1958). The ways in which technology enables and constrains how people behave also discussed in relation to politics of technological artefacts (Winner, 1980). Social media as platform-based businesses exemplify this well. Their complex sociotechnical machinery is constituted of specific tasks assigned to their users and a specific form of technology facilitating these tasks according to these platforms' business imperatives. This is accomplished by constructing data representations of users through structured platform features and by continuously tweaking and reshaping those features with analytics tools. This indicates that there is a cyclical

relationship between how people are represented through engineering devices, and, how these representation shape behaviour.

This relationship paves the way for the mutual calibration of platform and the data it produces about people in the course of building and optimising a social media platform. In this regard, it raises significant questions about the politics of sociotechnical arrangements like social media and the means they use to create data representation of individuals (Gillespie, 2010; Puschmann and Burgess, 2013; Ruppert, Isin and Bigo, 2017). In a recent commentary Ruppert, Isin and Bigo, 2017) raised such questions about what they refer to as 'data politics'. They define this term as involvement of data and related data practices in the reconfiguration of power/knowledge nexus (Foucault, 2007) through the new forms of data accumulation about people and their behaviour. They argue that the producers of new forms of data about people exercise power over them because "*data enacts that which it represents"* (Ruppert, Isin and Bigo 2017, p.1): *"To collect, store, retrieve, analyse, and present data through various methods means to bring those objects and subjects that data speaks of into being. Data sciences such as statistics, probability, and analytics have emerged not because they have merely quenched our curiosities but because these sciences have been useful for the objects and subjects they have brought into being for the purposes of governing and/or profit."*

This emphasis on data's power to shape and constitute what it represents is closely related to the notion of performativity. In the following sections I will elaborate on this by discussing the implications of creating data representations of people in social media and how this is related to shaping their behaviours in various ways.

8.3. Representing Users in Engineering Devices

This study unpacked the relationship between the *frontend* and *backend* of social media in light of the concept of platformization. This is done to gain a better understanding of social media's matchmaking operations. This theoretical approach helped to establish the logic underpinning how individuals' tastes are digitized based on certain assumptions about 'what is taste' and 'how it can be chased and rendered computable'. As empirically illustrated, the *supposed explanatory power* of pre-formatted user actions (i.e. like, tag, share etc.) constitutes the basis for the ways in which it is taken as a variable to define music taste. This also involves how much weight this variable is given in the definition of taste. As

demonstrated in the analysis, each user-action is given different weight when computing users' music taste. Furthermore, it is also demonstrated in the analysis how this is combined with the perceived 'importance of an action' by the developers for the operations of a social media platform (Bucher, 2012). The action of 'tagging', for instance, is given the highest weight by developers when computing music tastes. Because this is a *desired form of behaviour* that developers expect from their users. This is due to the platforms' reliance on users for categorizing content using tags (Webster et al., 2016), a notion also known as 'folksonomy' (Lamere, 2008).

This illustrates how *developers' assumptions* shape the criteria used to define and compute aspects of people using social media platforms (Webster et al., 2016; Bucher, 2012), such as their taste. This is a significant empirical contribution advancing knowledge on social media's knowledge production mechanisms. This contribution is very important particularly given that these aspects are significantly overlooked in the existing state of the literature. Even if they are recognized it is very difficult to empirically demonstrate these aspects since owners and developers of these platforms are reluctant to share the source of their competitive power. The crucial point here is that representing people in the form of data through engineering devices is not a straightforward process. An individual's taste, which is a highly ephemeral and abstract aspect of one's identity, is turned into data as a result of a set of assumptions regarding what constitutes taste and how users behave. More importantly, these assumptions become intertwined with the expectations regarding how users *should* behave in ways aligned with the business objectives of the platform owners and system developers. These expectations are not free from consequences. They shape the prescriptive design of the platform and shape user actions in particular directions to produce desired form of behaviours. This illustrates how platforms govern and control user behaviours, signifying how they are able to exercise power over people. They do so in highly subtle and not easily recognizable ways. For this reason they are seldom questioned. This signifies how the social media's distinct knowledge production mechanisms are linked to the new forms of power. In the following sections, I will elaborate on different ways in which social media exercise power over users. In doing so, I will also discuss why social media platforms should be distinguished as specific types of platforms that have specific forms of power to control and govern users' behaviours and shape their preferences.

8.3.1. Importance of Data Over Algorithms

This study demonstrated that the relationship between the front-end and back-end of social media is not fixed. On the contrary, it is shown how it is dynamically changing in accordance with the concerns related to how to scale the platform business. Platform design changes based on the assumptions and expectations about the desired user behaviours, and this occurs with the cyclic use of data-based means such as analytics. These tools are deliberately used to create certain forms of behaviors. Developers do this by continuously tweaking the platform features through a series of hypothesis testing. They verify or falsify their hypotheses about what motivates users to perform specific behaviours. These hypotheses may take the form of suggesting auto-tags while creating playlists to encourage users to categorize content, or it can be using an up-voting system to rank users' curations to increase their intrinsic motivation to further participate. As demonstrated in the analysis, the main reason behind these two hypotheses testing was often under-recognised, yet a crucial fact that users are the main drivers of *content* and *data production* in social media platforms (Couldry and Kallinikos, 2017; Van Dijck, 2013). Hence they are the sources of the scalability and growth of these businesses (Parker et al., 2016).

What is important to notice here is that these business concerns are significantly implicated in the ways in which tastes profiles are created and feed back to users' preferences. These business concerns are materialized in the form of changes in the platform features, which script and encode user behaviors (Alaimo and Kallinikos, 2016, 2017). These points illustrate that 'raw data' is *indeed* an oxymoron (Gitelman, 2013) as data are variously 'cooked' in the processes of their collection and use. This study demonstrated this by highlighting some of the ways in which data are carefully crafted and cooked in accordance with the interests of platform owners in the course readying data for algorithmic operations (Gillespie, 2014). These findings validate the previous views that suggest the cultural implications of social media's algorithm cannot be considered by divorcing them from the larger sociotechnical context within which they operate (Alaimo and Kallinikos, 2017; Couldry and Kallinikos, 2017; Gillespie 2010, 2014, 2018).

Algorithms are usually subsumed under the more generic term recommender systems that are considered to be operating as 'tastemakers' (Santini, 2018) and shaping cultural taste (Webster et al., 2016; Beer 2013; Morris 2015a). Although constituting a highly crucial component of the production of personalised recommendations, algorithms alone cannot

capture the complexity of the wider socio-technical configuration through which data representations of taste are produced to feed these recommender algorithms. This is especially important considering that algorithms without data are just mathematical fictions (Kallinikos and Constantiou, 2015) that are "*inert, meaningless machines, until paired with databases upon which to function*" (Gillespie 2014, p.169). According to Alaimo (2014), algorithmic operations take place in the final stage of a complicated set of data operations built on coding user actions and make them ready for algorithmic processing. Kallinikos and Constantiou (2015) share this view. They argue that, "*the conception of algorithms and the function they fulfil in the various contexts of computing are easier to grasp than the invisible and more diffuse work of the structures and arrangements through which data is produced and made available for algorithmic calculation. The cardinal importance of data structures emerges against the background of the painstaking work their development requires. It is on the basis of such elaborate datawork that data is delivered in formats amenable to algorithmic calculation and, more generally, machine processing*". (p.73)

These views are in line with Gillespie's (2014) emphasis on why sociological studies on algorithms need to take into account the role of data. Similarly, Van Dijck (2013) stresses the importance of understanding the work of algorithms in relation to other elements, namely *data* and *metadata*, *interfaces*, *protocols*, and *defaults*, which according to her are the core co-constituents of the distinct sociotechnical configuration of social media: *"The providers of software, (sometimes) hardware, and services that help code social activities into a computational architecture; they process (meta)data through algorithms and formatted protocols before presenting their interpreted logic in the form of user-friendly interfaces with default settings that reflect the platform owner's strategic choices".* (Van Dijck, 2013: p.29).

This case study on a social music platform confirms these views. In this regard, it makes an important empirical contribution to knowledge by demystifying the role and power of algorithms. This is accomplished by providing evidence showing how data is made ready for algorithms. It is demonstrated in detail how the platform system encodes listeners' behaviours in the front-end through the distinct design logic that structures user actions on the user-interface level. This allows aggregating user actions to turn them computable data. Thus, user dispositions are digitised and rendered data to construct user profiles. It is only at the *final stage*, as Alaimo (2014) underlines, the algorithms process this data to compute users' tastes, and construct personalised recommendations matching users tastes. This is why,

according to Kallinikos and Constantiou (2015), social media and their personalised recommendations cannot be explained by merely focusing attention on their algorithms. Instead, in Wenger's (1997) words, they should be considered as massive 'interaction machines' whose functioning depends more on the specific type of interaction data they produce rather than their algorithms. This is why algorithms, Kallinikos and Constantiou (2015) argue that they need to be distinguished from the broader category of software automation: "*There is much more to programming and software engineering than the development and design of algorithms…. Data structures represent systematic and stratified arrangements through which data is captured and ordered, while algorithms provide a fundamental means through which such data is computed*." (p.73).

These views illustrate the importance of understanding the work of recommender systems' algorithms in relation to the larger sociotechnical arrangements such as social media. By exposing social media's distinct logic reliant on production of interactive data this thesis showed the significance of the data-work preceding the operations of algorithms. In this regard it makes an empirical contribution to the literature by providing evidence for why algorithms cannot be considered as stand-alone actors responsible for shaping taste.

8.3.2. Social Media's Reliance on Users

It is important to recognise that social media are just an instance of the broader category of online platforms, which provide cultural services utilising personalised recommendations. These may be streaming services like Netflix and Spotify and/or retail stores like Amazons, which sell cultural goods like music, books, films (Beer, 2013; Morris, 2015a; van Dijck, 2013; van Dijck et al., 2018; Webster et al., 2016). However, of particular importance for this study are how social media platforms are reliant on the work of their users for content production. I will now elaborate on why this reliance matters.

Platforms such as YouTube and Last.fm provide cultural services based on *user-generated* videos or playlists rather than offering content owned or distributed by the platform. For instance, Netflix and Spotify hold the right to distribute cultural content such as music and films that reproduced by professionals. These kinds of content distributed/streamed via online platforms based on the licencing agreements made with the recording labels and production companies. Some of these platforms may have features that are typically found in social media platforms (i.e. Spotify allows users to create and share their playlists). However, it is important to recognise that they do not depend solely on their users for curation and

categorisation content. Rather, these platforms rely on the experts they employ to curate and categorise the cultural goods they distribute.

Social media, on the other hand, are significantly reliant on the work of users (van Dijck, 2013; Webster et al., 2016). Users curate and categorise content typically streamed from elsewhere, if not produced by users. They are also reliant on users for data production (Couldry and Kallinikos, 2017). Previous studies demonstrated that social media produce data about their users not by depending on pre-existing life activities but rather by encoding online sociality in ways programmed by the distinct design of the platform. In this regard, Alaimo and Kallinikos (2017) argue that social media produce a distinct type of *'social data'*. Their matchmaking operations are hence considered to differ from other platforms'. This study empirically demonstrated this by showing how social media 'plat-form' cultural intermediation in ways distinctly different than other online platforms do. It is shown that the chosen platform's design logic allows steering its users' mundane behaviours by engineering their daily interaction with music to compute their music taste. This validates previous conceptualisation of social media as '*post-transactional spaces*' *(*Alaimo and Kallinikos' 2017), where facilitating interactions around transactional activities such as buying, selling, booking or streaming etc. are *less important* than building a 'user-labour reliant' space to craft their taste profiles out of the daily trivia they engage. Thus shown how the users and the platform share the curation work in social media.

8.3.3. Conceptualising Social Media as Sociotechnical Cultural Intermediaries

This research contributes to existing knowledge by empirically illustrating the sociotechnical makeup of social media, which includes, but not limited to: a distinct design logic underpinning platform-based business models; The work of users whose role is strictly designed according to this platform logic; unique prescriptive design choices standardising and shaping user-platform participation on the interface level in ways generating data for back-end data operations; data about user behaviours and metadata describing content; algorithms and mathematical models used to process data; assumptions and expectations of developers designing the platform system; data-based means such as analytics and accompanying data analysis techniques; a unified and standardised data infrastructure built in accordance with the requirements of other services specified in their *Application Programming Interfaces* (APIs) for aggregating data and metadata to profile users and content.

The last points is related to how complementary to the previously discussed points, this study confirms previous work highlighting the importance of APIs in social media's configuration and the ways in which they afford platformization (Helmond, 2015; Plantin et al., 2018). It is shown that social media do not own cultural content but rather depend on other services for this (i.e. free services such as YouTube, SoundCloud or Streaming Services). It is also demonstrated that they also aggregate metadata from other platforms (i.e. EchoNest, Last.fm). This is another important empirical contribution of the thesis. All of these operations, as this study illustrated, are done in accordance with the requirements specified in these services APIs. It is discussed in the analysis how these APIs make it possible for these services to connect with each other. Moreover, they are involved in shaping individual platforms' data through the requirements rendering data interoperable within an ecosystem of connected platforms. van Dijck (2013) refers to this as 'the ecosystem of connective media': *"If the aim is to understand how, in the intervening period, online sociality evolved, it is not enough to study individual platforms; rather, we need to apprehend how they coevolved in a larger context of interpenetrating platforms and to dissect the cultural logic undergirding this process. Therefore, I propose to look at distinct platforms as if they were microsystems. All platforms combined constitute what I call the ecosystem of connective media—a system that nourishes and, in turn, is nourished by social and cultural norms that simultaneously evolve in our everyday world. Each microsystem is sensitive to changes in other parts of the ecosystem: if Facebook changes its interface settings, Google reacts by tweaking its artillery of platforms; if participation in Wikipedia should wane, Google's algorithmic remedies could work wonders.*" (p.21).

Aforementioned non-exhaustive list of elements of social media's configuration extends previous views on the role of recommendation systems as sociotechnical cultural intermediaries constituted of the intermingling of human and technological actors (Karakayali et al., 2017; Lange 2016; Santini, 2018; Webster et al., 2016). For instance, Morris (2015a: p. 456) argues that recommender systems gain authority as tastemakers due to the '... cultural knowledge of those creating the databases and algorithms' (p. 456). Similarly, Webster et al. (2016) conceptualise music recommenders systems as computational systems that exercise power over taste formation. Their work extends Morris' views by describing in more detail how changes in the process of cultural intermediation materialise and who and what have contributed to this. In doing so they focus on the role of 'human' and 'technological' factors and conceptualise music recommender systems as instances of "*complex sociotechnical*

systems made up of people, technologies, knowledge, data, algorithms and other heterogeneous actors" (Webster et al., 2016: 13). Drawing on Bourdieu's notion of habitus, their work highlights how the categorisation of cultural content is shaped by the cultural capital of users while and the behaviours of the recommendation algorithms are regulated by engineers' decisions, which are guided by their perception and appreciation of culture.

Webster et al.'s (2016) work improve our understanding regarding contemporary forms cultural intermediation by conceptualising recommender systems as co-creation of a range of human and technological factors, such as the labour of users; recommendation algorithms; and the designers and engineers' cultural assumptions. Confirming Webster et al.'s (2016) insights, it is shown that users and developers of the systems are indeed essential parts of the transformation of cultural intermediation. This empirical contribution is made in two ways: First, evidence revealed that users' work is fundamental to social media's functioning. They are the drivers of content production and categorisation. In addition, it is also demonstrated in detail how participation of users is structured by the prescriptive design of the platform (van Dijck, 2013; Alaimo and Kallinikos, 2016, 2017; Couldry and Kallinikos, 2017) and the data-based means used to optimise the platform features. Secondly, these aspects are also discussed in relation to platform owners'/developers' desires and expectations about user behaviour and the kind of data generated on the platform. It is shown that, user behaviours and resultant behavioural data are structured and shaped according to the requirements of platform business and their distinct business models that rely on scaling the platform through network effects (Nieborg and Poell, 2018; Parker et al., 2016).

The notion of platformization helped to better account for this structuring in ways that allow elaborating on social media's cultural implications. It highlighted how business imperatives of social media are implicated in shaping the role of user as well as how knowledge about their identities is produced (Couldry and Kallinikos, 2017; Gillespie, 2010; Helmond 2015; Langlois and Elmer, 2013; Van Dijck 2013, 2014; Van Dijck et al., 2018). This illustrates how, in Van Dijck's (2013) words, the platform developers strive to orchestrate the interaction between the users and the platform in line with their commercial interests. In this regard, this study has taken previous views a step further by providing context for sociotechnical cultural intermediation.

This is a significant empirical contribution. It also paved the way for one of the theoretical contributions of this thesis. Building on extending former views on the interplay of human

and technological actors, users and algorithms are conceptualised in this study as part of larger sociotechnical structures such as social media, which utilise these recommendation systems according to their platform-based business logic. The notion of platformization helped to do this by shedding light on how users' and algorithms' roles are allocated in social media, sociotechnical configuration of which exemplifies a wider transformation that is altering the economic, social and cultural domains (Parker et al, 2016; Constantinides, Henfridsson & Parker, 2018; Nieborg and Poell, 2018; Langley and Leyshon, 2017; van Dijck 2013; Alaimo and Kallinikos 2016, 2017).

Consequently, this thesis conceptualises social media as *sociotechnical cultural intermediaries***.** It explains how they 'plat-form' cultural intermediation by internalising the curation, guidance and measurement processes, which in this process merge, get layered and shared out by the platform system and the users. While platforms actively curate, choose, and select content (Gillespie, 2018) users do so in ways mediated by the platform according to its design logic (van Dijck, 2013). In this regard, this conceptualisation is in line with Van Dijck's (2013, p.36) theorisation of *user, technology, and content* as 'three fibres' that social media platforms 'weave together in order to create the fabric of online sociality', which they themselves engineer. As discussed earlier in this chapter, laying bare how social media turn cultural tastes into data and make them ready for algorithms considerably helped in extending these views as these processes are often overlooked or taken for granted. In this regard, this thesis makes a contribution to research on social media and cultural intermediaries by addressing this gap in explaining how knowledge about culture is produced and feedback to users and how this signifies a new form of power dynamic between the institutional actors like social media businesses and ordinary individuals like their users.

Van Dijck (2013) argues that social media operate as 'performative infrastructures' underpinning a new form of sociality co-produced by machines and humans. This notion of *performativity* points out here the significance of *self-fulfilling prophecy effects* created by quantification, measurement, and data production practices (Espeland and Stevens, 2008). These methods considered having the power to shape the phenomenon they explain. For instance, previous studies used the concept of performativity to explain how language has the power to shape what it represents (Butler, 1997) or how economic models shape the economic activity they purport to describe (Callon, 1998; MacKenzie, 2006). Similar performative mechanisms are found in the functioning of social media. These are, according

to Alaimo and Kallinikos (2017), are closely related with how social media create a "dynamic regime of quantified interaction between user *data* and user *behaviour*, whereby data generated by users are processed and fed back to them variously shaping their behaviour" (p. 176 emphasis added). This is considered to create a vicious circle like the relation between the users' data and their behaviour. Implications of this will be discussed in the next sections by elaborating on how user behaviours and the platform features mutually calibrate each other and how this is implicated in making up users' taste.

8.4. Mutual Calibration of Platform And Behaviour

As discussed earlier, social media's functioning significantly relies on user activity for content production/curation, categorization and equally importantly for behavioral data production (Couldry and Kallinikos, 2017; Van Dijck, 2013, 2014; Van Dijck et al., 2018; Langlois and Elmer, 2013). This study provided evidence that extends these views. It demonstrated in detail how a stream of data about user activity is also needed to fuel *analytics* tools used to *optimize* the platform in line with the business imperatives. This is another noteworthy contribution this thesis makes to this literature. Findings revealed that monitoring and orchestrating user-platform participation with analytics tools recursively shape the platform features and user behaviours. Moreover, it is highlighted how this is a result of platform owners'/developers' expectations from the users for the desired form of behaviours and the efforts put in creating them. In this section, I will elaborate on these findings and discuss the implications in light of the previous views shedding light on some of the ways in which such expectations become self-fulfilling prophecies.

8.4.1. Feedback, Reactivity and Self-fulfilling Prophecy Effect

Previous research on social media invited attention on some of the ways in which data produced about people by social media variously *feed back* to them *influence* their *behaviour* (Alaimo and Kallinikos, 2017; Bucher, 2012; Gerlitz and Lury, 2014). This relationship between the platform design, behaviour, and data exemplifies what Beer (2015) refers to as the 'productivity of measures'. Previous works stress that, metrics on social media do not merely measure participation but rather they become 'participative metrics' co-producing what they attempt to measure (Gerlitz and Lury, 2014; van Dijck and Powell, 2013). In other words, measures shape what is measured.

This study confirms these views by empirically demonstrating how the analytics tools measuring users behavior *do not* capture a *separate reality*. Rather, as Gerlitz and Lury (2014) emphasize, they are deliberately employed to modify the activity they themselves invite. To illustrate, this case study on a social music platform revealed that presumably *passive* activity is not a valued and a desired form of participation. The reason behind is that, passive forms of participation like listening only fuel the consumption side of the platform. However, the dominant view is that consumption *per se* cannot scale the platform without a satisfying amount of content (Parker et al., 2016). For that reason, it is argued that the consumption side needs to be constantly supplemented by the production side of the platform. It is shown how for this reason developers and analysts in the studied platform focused their attention on *motivating* and *encouraging* users for participation, preferable in the form of '*frequent'* and '*active'* contribution to data and content production. This is illustrated by demonstrating some of the strategies developed to this end. For instance, metricizing and quantifying user actions; notifying producers about consumers' reactions to their curations; and providing personalized recommendations are identified as utilised methods to create *self-reinforcing feedback mechanisms.* The aim behind this was to motivate users by creating intrinsic motivation to generate the desired form of behaviours on the platform.

This empirically validates how 'feedback loops' are widely used in social media as commonly utilised gamification techniques (Lampe, 2014; Pellikka, 2014). These are automated and self-regulating mechanisms, which illustrate the techniques that are deliberately used as a form of *control* by shaping the *norms* of behaviour (van Dijck, 2013). These aspects, according to van Dijck (2013), are closely related with how platform owners strive to *construct* a 'particular kind of user' that is *active* and *engaged*. As illustrated in this study, 'active' user participation is accomplished by architecting specific governance strategies within the platform design. These are designed in the form of feedback loops and carefully embedded in the platform's functioning. The analysts, who use the analytics tools to see if they produce the desired form of outcomes, continuously assess the effectiveness of these strategies. For example, they analyze the 'conversion rates' to see whether there is a change in user behaviors towards the valued direction. As shown in the analysis, an instance of this was when they looked for an increase in the number of users completing the action of 'tagging' a playlist after introducing the automated tag suggestion feature.

Aforementioned examples illustrate some of the ways in which user participation is incentivized and become controllable in social media. These mechanisms are recognized in the literature as distinct strategies used to keep users engaged to strengthen the network effects (Parker et al., 2016). This study contributed to the literature by empirically demonstrating some of the ways in which these are architected in the platform design and optimised over time as a distinct form of *governance strategy*. This points to a striking aspect of the relationship between the development and optimization of a social media platform. Data analytics tools continuously inform the improvement of the platform design in ways that result in a mutual calibration of the platform and user behaviour. This is mainly because feedback loops rely on *reactivity*.

The notion of 'reactivity' explains the ways in which measures intervene in the social worlds that they depict. Espeland and Stevens (2008), assert that, "measures are reactive, they cause people to think and act differently" (p. 412). Because, they argue, numbers exert discipline on those they depict: "measures that may have initially been designed to describe behaviour can easily be used to *judge* and *control* it" (p.414 emphasis added). For instance, if we go back to the analysis of the evolution of the studied platform, we find that the studied platform saw more value in involving its users in the content production. In their words, this was more *'scalable*' than doing it by merely relying on the algorithm to aggregate content. For this reason, the desired behaviour became 'active participation in content production'. In order to shape the user behaviour accordingly, the platform used several strategies such as visibly measuring user-producers' performance (i.e. their visible taste profile showing mastery level on certain genres; up-voting system ranking users; number of followers etc.). This way, developers' assumptions regarding how expected user behaviours can be created (i.e. how users can be motivated to produce more) become *'testable hypotheses'*. These hypotheses are then tested by frequently tweaking the design choices until the desired results are achieved. This denotes that, *predictions* regarding user behaviours may directly or indirectly *become true* if these behaviours are continuously monitored and channelled by modifying the design features using analytics.

In this regard, this study contributes to advancing knowledge on how social media platforms are prone to create self-fulfilling prophecy effects through the ways in which they exploit users' reactivity (Gerlitz and Lury, 2014). Findings of this research confirm how the forms of 'reactivity' and 'self-fulfilling prophecy', which are considered to be a problem in some forms

of measurement (e.g. Espeland and Sauder, 2007), are actually an intentionally created effect in social media. As Gerlitz and Lury (2014) assert, social media metrics (such as number of 'likes', follower etc.) are purposefully presented to users in ways that would develop *a state of reactive tendency*. These insights are complemented by Grosser (2014) who emphasises that these metrics are specifically employed design choices and strategically made visible to users based on their perceived ability to increase or decrease user participation.

Scholars who studied software systems previously provided similar insights about the performativity and self-fulfilling prophecy effects (Mager, 2012; Kitchen and Dodge 2011; MacKenzie, 2006). For instance, Mager (2012) explained how software becomes a part of social action *not just* through what it does *but also* based on how individuals and organisations respond to what software does or how they imagine what it does. In turn, Mager (2012) adds, the system developers tweak the software system itself in response to these reactions in order to maintain the control over the user and system behaviours. Thus, Mager (2012) argues, the software systems are shaped and reshaped in response to user reactions as well as developers' strategies to prevent certain forms of behaviours and to produce the desired ones.

This illustrates how social media platforms, which are essentially software systems, and user behaviours recursively shape each other. The central concern here, according to MacKenzie (2006), is the issues of ordering, control, and power to shape behaviour. He argues that software systems have the capacity to order in ways that normalises this ordering power. They are invisible structural forces shaping everyday life in subtle yet highly significant ways. Because, they have the capacity to shape possibilities, limit agency, and affect the social world. They do so by "*selecting and* reinforcing one ordering at the expense of others" (Mackenzie, 2006, p.44). The crucial point here, according to MacKenzie (2006), is to recognise how agency is contested as software affects what can be said and done by naturalizing particular orders and animates certain movements.

In the same vein, Beer (2013) argues that the design of software systems is shaped by the judgment of the developer's understanding of the likely outcome, which in turn shape users' behaviours. He argues that, this may happen as users' behaviours and actions change while they try to *predict* and *manipulate* the outcomes of algorithmic recommendations. Bucher's (2012) work Facebook illustrates these points well. She highlights how Facebook's newsfeed operate as a filtering mechanism making decisions about whose activity will be placed on top

of others and thus become more visible. This, as Bucher (2012) stresses, paves the way for certain type of behaviours as users try to play with the platform system with the aim of being more visible. Bucher's work is another example of how social media platforms are designed to exploit users' reactive tendencies (Gerlitz and Lury, 2014). They exercise power over user behaviours by treating them with invisibility, while rendering visibility a form of *reward* users gain when they exhibit a certain form of behaviours. This again illustrates, social media strives to create 'a particular type of user' (van Dijck, 2013)

This is closely related with the previously discussed notion of performativity. This concept discussed so far in relation to how behaviours of people are recursively shaped by the platform through the means employed to monitor user-platform participation. Another prominent manifestation of the performativity of this vicious circle like the relation between the platform and user behaviour is the construction of personalization recommendations. In the next section I will elaborate on how user behaviours are mutually shaped by the platform system and, how the resultant data representing users taste feeds back to users and shape their preferences. I will then discuss the implications of this process in relation to the notion of tastemaking.

8.5. Matchmakers or Tastemakers?

As discussed earlier, the term 'performative' is defined as the capacity of quantification and measurement to change the social reality they represent (Espeland and Stevens 2008). From this viewpoint, I argued how social media's data-based means used to monitor, measure user behaviours are prone to shape these very behaviours and engineer them. These aspects were discussed in light of previous views pointing out how social media are creating 'a particular type of user' in ways of fulfilling their business interests (van Dijck, 2013)

Previously discussed mutual calibration of the *platform* and *behaviour* also shapes the kind of *data* social media platforms produce about users. Studies have pointed out how this resultant data feeds back to users in various ways, most notably through the personalised recommendations (Alaimo and Kallinikos, 2017). This study advanced this insight regarding how social media constantly dissect and reconfigure user-platform participation by data-based recommendations. It is also shown that these recommendations are produced through the ways in which user interactions are organized, recorded, and measured. This indicates how, "*the patterns of sociality that emerge online (with whom to interact, when and how) are*

shaped under conditions that reflect how social media platforms organize user participation and direct user attention via a series of computations. A range of scores and measures of aggregate user platform activity (e.g. similarity, popularity, trending) are routinely computed by social media and cycled back to users in the form of recommendations and personalized suggestions" (Alaimo and Kallinikos 2016, p.78).

These recommendations are presented to each individual user in personalized ways based on the behavioural profile constructed for each user. As demonstrated in this case study every time a user takes new action, her taste profile is updated. Accordingly, what she sees as suggested content also changes. Cases when the user takes up and acts upon such recommendations (i.e. start listening to a playlist recommended to her) can be considered as a form of shaping her behaviour *once again* as without this type of mechanism she might have listened to a different content. This corroborates with what Beer (2013) describes as the ways in which 'systems of measurement' *feed into* and become active in shaping and constituting taste. He elaborated these implications by drawing on the personalised recommendations provided by the movie streaming service Netflix. He noted: "*it is likely that the films recommended to you by these algorithmic processes are likely to become the films you watch or that you are likely to want to watch*" (Beer, 2013 p.81).

Building on these views and by drawing on the empirical findings of this study it can be concluded that chasing something as abstract as taste (i.e. listeners' music taste) is prone to operate in *performative* ways. To explain, this case study illustrated how the assumptions about the correspondences between '*persons*' and '*things*' are built into the data-structures underpinning matchmaking operations of social media platforms (Alaimo and Kallinikos, 2017). In turn, the same assumptions shape what individuals encounter on the screen because they shape how the platform selectively presents consumers with a set of user-curations in the form of personalised recommendations. This echoes with what Lash (2007) refers to how things come to 'find us'. He argues that we do not find content by explicitly searching for it anymore. Instead data and content find us as a result of particular decisions embedded in the ways in which data is filtered.

This is closely related with how social media platforms implicitly 'directs attention' (Alaimo and Kallinikos, 2016; Beer, 2013) towards particular products and make them 'visible' (Bucher, 2012) based on certain assumptions about the 'relevance' (Gillespie, 2014). Other content, which identified based on the same assumptions as less relevant, get left out and

rendered invisible. "*Drawing attention to a desired focal point, thus obscuring other possibilities*", in this way, as Beer (2013) asserts, is a 'form of power' exercised over how encounters are directed in ways implicated in how users' decision making processes. Shaping encounters as such by establishing the relationship between *'persons*' and *'things*' is considered to be a performative mechanism: *"One of the effects of abstracting the world into software algorithms and data models, and rendering aspects of the world as capta, which are then used as the basis for software to do work in the world, is that the world starts to structure itself in the image of the capta and the code - a self-fulfilling, recursive relationship develops"* (Kitchen and Dodge 2011, p.41). In this vein, Beer (2013) argues that data-based predictions embedded in personalised recommendations have the potential to become a reality as these predictions become the basis for actionable decisions: "*one of the key powers of algorithms is in the predictions of the future that they are used to make. These predictions, as we have suggested, become realities from which decisions are made."* (P.87).

Making these data-based predictions are closely related with classifying and ordering people, which ultimately shape how they are treated and what type of response they receive. From this perspective, Cheney-Lippold (2011) assert that sorting and categorising people has the power to order and control: "*The power here is in defining and differentiating us through data, which, in turn, creates normalising behaviours and identities* (Cheney-Lippold, 2011: 177). These issues are closely related with the broader debates in the literature concerned with the implications of categorisation, quantification, and measurement. In the remaining of this section I will elaborate on this in particular with respect to Ian Hacking's (2002) views on *kind*-making.

Ian Hacking previously stressed how measurement and categorization processes result in 'looping effects' (1995). This term defined by Hacking as the way in which classification may *interact* with the people classified and result in changing them. In this regard, he argues that means of categorization used in investigating behaviours result in 'making up people' (Hacking 1986, 2007). Hacking calls these means 'engines for making up people' (Hacking 2007, p.3). He describes the statistical analysis of classes of people as a fundamental instance of such engines. Similar engines for making up people such as censuses have long been used by states and other institutions to produce knowledge about subjects to control their behaviour (Desrosières 1998; Porter, 1995). In this vein, drawing on Foucault's (2007) ideas about the relations between power and knowledge, Ruppert, Isin and Bigo (2017)

conceptualise ‘data’ not as a mere representation of facts, but as a means to exercise power. According to their view, data does not just represent, it enacts that which they represent often times in ways that serve the interests of those who produce it. In this regard, they consider data having a ‘performative power’: “*rather than settled in a database or archive, data has a potential force that can be realized in myriad ways through its uptake and deployments*” (p.2).

This study advances scholarship by providing evidence for how these engines for making up people have taken new forms. Previously discussed data-means involved in how social media investigate people’s behaviours and define them *qua* users (Alaimo and Kallinikos, 2017; Couldry and Kallinikos, 2017) can be considered as instances of the new forms engines used to produce knowledge about people. As empirically demonstrated, social media categorize individuals by aggregating their behaviours through the ‘tags' describing the ‘objects' of their ‘actions'. These categories then inform the ways in which these individuals and their taste are defined and used to construct personalized suggestions. Consequently, as people take up these recommendations, these categories get naturalized over time, and thus, a new form of category-naturalization occurs (Couldry and Kallinikos, 2017).

How this category naturalization occurs is evident in previous studies on social media. For instance, Karakayali et al. (2017) found that Last.fm users intentionally modify their behaviours to play with and manipulate the recommendation outcomes in order to cultivate their music tastes. This illustrates what Mager (2012) describes how software shapes user behaviours based on how individuals imagine and respond to what software does. As discussed earlier, developers tweak the platform system in response to users’ reactions. Thus developers’ aim to maintain control over the user lead to recursive shaping them as users interact with the ways in which platform system defined them.

This is in line with Hacking's (1999, 2007) description of making up people through the looping effects. This mechanism renders the objects of categorization ‘moving targets'. Because investigations about individuals *interact with* these people and eventually change them: “*since they are changed, they are not quite the same kind of people as before. The target has moved. That is the looping effect. Sometimes our sciences create kinds of people that in a certain sense did not exist before. That is making up people*” (Hacking, 2007: p.2). Karakayali et al.’s (2017) study on the social music platform Last.fm illustrates this mechanism well. Drawing on the interviews conducted by the Last.fm users, they revealed

that platforms like Last.fm changes users perceptions. As a result, users transform themselves. Karakayali et al. (2017) use the term 'askesis' describe the changes in people regarding "how one lives and conducts oneself through self-disciplinary labour" (p.7), in ways leading to a self-transformation: "*Last.fm, then, functions as a technology of the self, not by urging users to interiorize a definite musical 'content', but by inviting them to modify their listening habits through recommendations. User comments indicate that in the absence of the recommender system, one might enjoy music without much concern about, say, how often one listens to the same song or how diverse the artists one listens to are. It is, in other words, by encouraging a change in this 'careless' conduct of the user as a music listening subject that the recommender system plays a role in self-transformation. This, moreover, implies an askesis because 'music taste' is now perceived as the outcome of an activity – besides enjoying music one also listens to it with a view to improving one's 'music taste'*." (Karakayali et al., 2017: p.13).

8.5.1. Rendering Individuals Unstable Data Objects

As mentioned earlier, Hacking (2007) describes people who are subjected to classifications as 'moving targets'. Because investigations into their habits interact with them and eventually change their behaviour. As a result, the knowledge about them also changes: "*new knowledge about 'the criminal' or 'the homosexual' becomes known to the people classified, changes the way these individuals behave, and loops back to force changes in the classifications and knowledge about them"* (Hacking, 1999: p. 105). This is discussed in relation to how classifying people can easily take the form of a *self-fulfilling prophecy*. As discussed above, one of the most prominent examples of how measures interact with the people they measure is social media's personalized recommendations. As discussed in the analysis, the studied platform defines users based on the data generated through *encoding* their heavily *prescribed actions*. Encoded data go through several rounds of normalization, aggregation, scoring, and computation. Thus individuals' tastes are re-constructed in the database in the form of data representations comprised of *top-ranking user-tags scores.* This illustrates how social media reconstructs individuals *qua* users in the database as 'new social entities' (Alaimo and Kallinikos, 2017). What is noteworthy to recognize here is the *unstable ontology* of these entities, definition of which *contingently changes,* since these data representations of individuals are continuously updated every time a user takes a new action.

In this regard, this study advances scholarly views by empirically showing that people using social media are *indeed* 'moving targets'. Investigations into their habits result in dynamic forms of knowledge regarding 'who they are,' which in turn, *interact* with them. Because, the platform acts upon this information while guiding them for the discovery of new content through presenting them personalized recommendations in particular order. As Karakayali et al.'s (2017) work reveals, users often take up these recommendations either by following or manipulating them. In turn, how users react to these recommendations is again encoded and used to update their data definitions. In this regard, this study makes an important contribution to scholarship by revealing how investigations about individuals in social media, such as measuring their taste, interact with them. As a result, people's data representations are recursively shaped and individuals rendered 'unstable data objects'.

8.5.2. The Role of Users

Another empirical contribution of this thesis is providing evidence regarding the fundamental role users play in the sociotechnical configuration of social media. This demonstrated by highlighting the importance of how content production is an essential concern for the platform, and, how this concern is implicated in defining and shaping users role in this complex arrangement. Sharing content is the key means to connect people by helping them to have a global stage for public viewing (van Dijck, 2013). In the studied case this was in the form showcasing music taste. Following Parker et al.'s (2016) ideas it is shown that the user-generated content constitutes the 'value units' through which the platform facilitates user interactions. It constitutes a crucial aspect of the platform's 'core interaction', which conceptualised in this study as the building block of a social media's formation as matchmakers. These aspects are elaborated earlier in this chapter as part of the discussion on how social media's operate as distinct kinds of sociotechnical cultural intermediaries due to their reliance on users for content and data production. This reliance is closely related with how social media shape their governance strategies to ensure the standardisation of content and uniform deliverance (van Dijck, 2013).

In the studied case, an example of these regulatory features was the mandatory steps to be taken while creating a playlist. Users were asked to follow a set of instructions, which force them to obey the platform rules. Only after following these rule users were able to share their music taste with other users. The same regulatory-means also used to ensure to produce knowledge about users by rendering them as 'data objects', which fuel the platform's

matchmaking operations. It is also shown that users interact with their data representations and become moving targets changed by the data means used to describe them.

This, however, does not mean that people cannot intervene in the ways which engines of social media constitute them as 'data objects'. Historically, there have been many instances where people challenged the categories used to define them. To illustrate, Ruppert, Isin and Bigo (2017) describe how citizens subverted and transformed census categories such as race, ethnicity, and gender etc., by calling them into question. Similarly, studies on social media showed how users interact with categories to deliberately manipulate the outcomes (Bucher, 2012; Karakayali et al., 2017; van Dijck, 2013). While this aspect is largely discussed with respect to the performative implications of rendering user data objects, it is noteworthy to recognise that decoding the system's logic and playing with the features indicate that users are not passive recipients of technology and they have agency.

User agency in the context of social media is a contested phenomenon. On the one hand, it is recognised that the majority of social media users are not aware of the mechanisms through which their behaviours and choices are recursively shaped, as they work "*without the knowledge of those upon whom they are taking an effect*." (Beer, 2009: p. 990). On the other hand, users are also "neither 'dupes' of technology nor they are uncritical adopters" of it (Van Dijck 2013, p.32). As they are subjected to categorisations they develop certain tactics to alter them. According to van Dijck (2013, p.6), this is "*precisely what happened with the development of social media platforms*" as "*users 'negotiate' whether and how to appropriate them in their quotidian habits*."

These points indicate that users play crucial roles in social media's functioning and their role is carefully crafted in accordance with the 'platform logic' (Parker et al., 2016). This study provided empirical evidence supporting previous views on how this setup shape user behaviours and the knowledge produced about them. However, the researcher also recognise that users have the agency to *subvert* or *resist* these mechanisms by *tinkering* (van Dijck, 2013) with the platform features to manipulate the outcome of its data operations. However, as discussed earlier, the use of analytics tools help platform developers to track these user *negotiations* and thus they tweak the platform system to maintain control over the user and system behaviour. In this regard, mutual calibration of social media platforms and user

behaviours involves both the tactics of users to challenge their assumed the role as well as the hidden data work they get subjected to.

8.5.3. Social Media's Engines for 'Making Up Taste'

In the previous sections, I discussed in detail how knowledge social media produce about users interact with and change them. In light of this discussion, this thesis proposes the concept of '*social media's engines'* for *'making up taste'*. This is to conceptualize the loop-like relation between the data and the ways in which this data is generated and feed back to users in social media. Given the ways in which taste is made known is considered as a 'new dimension' shaping taste (Wright, 2015), the sociotechnical means involved producing knowledge about taste can be conceived as social media's engines for making up the taste. This corroborates with Beer's (2013: p.81) views describing how personalised recommendations: have "*the capacity and potential to make*" (Beer, 2013: p. 99). Wright (2015) shares his view. He asserts that knowledge about taste, produced by measuring people's behaviours feed back to them and shapes their decisions. This according to Lash (2007) is how data make the world; it folds back into the social world and begins to not only capture it but also 'constitutes it'. This in his view signifies that data-based arrangements like social media exercise a new form of power. Lash (2007) refers to this as a form of 'post-hegemonic power' that is not driven by ideology and does not act on subjects from above. Rather it is ingrained for of power that operates by shaping and constituting the social world.

8.6. Conclusion

This chapter discussed the key contributions of this research by highlighting how it builds on, validates, and advances previous theoretical accounts that informed this research. In light of the empirical analysis of the platformization of cultural intermediation, the 'plat-formed' versions of curation, guidance, and measurement are discussed in terms of how they merge, layered and shared out by social media platforms and their users in the course of their internalisation by the platform. Thus, social media platforms are conceptualised as sociotechnical cultural intermediaries and their unique aspects are discussed at length. The main contributions of the research are discussed around the themes of representing users' taste through engineering devices, mutual calibration of the platform and user behaviours', and how these may be implicated in shaping cultural tastes. As part of this discussion, a new concept is developed to explain the complex mechanism underpinning the ways in which social media, as sociotechnical cultural intermediaries, participate in tastemaking. The term

'social media's engines' for ‘making up taste' is proposed to point out the performative aspects of the data-based means social media use to investigate user behaviours and define their tastes. Thus this thesis advanced scholarship by making various empirical and theoretical contributions to the literature concerned with the social and cultural implications of social media platforms, by conceptualising them not only as matchmakers but also contemporary tastemakers shaping their users’ tastes.

9. Conclusion

This chapter concludes the dissertation by summarizing the main points. It discusses the contributions and highlights some of the ways in which findings of the study can guide future research. Finally, I highlight possible routes for future research by reflecting on the limitations of this study.

9.1. Overview of the Thesis

This research examined how social media transform cultural intermediation in the course of building and optimizing a digital platform. The research started as an inquiry on the transformation of traditional cultural intermediation processes by social media. It problematized the current views on the literature that focus attention on the involvement of users and/or the recommender algorithms to explain this transformation. An in-depth case study on a social music discovery platform is conducted to explain how social media alter this process beyond the role of the users and the recommender algorithms. Preliminary theoretical ideas are refined in light of the empirical findings and consequently social media platforms are conceptualized as 'sociotechnical cultural intermediaries'. The concept of platformization is used to explain the distinct ways in which social media internalise the traditional cultural intermediation processes. Empirical case study revealed that, these processes get layered and are shared out by the users as well as the platform system. It is also revealed that that these processes are underpinned by data operations through which users' taste rendered data and readied for algorithmic recommendations provided to users by the platform. Thus, the role of users and the recommender algorithms are conceptualized as part of the complex sociotechnical machinery of social media. Implications of these findings discussed in relation to larger debates in the literature and linked to the former views focused on the changing aspects of tastemaking. Consequently, a new concept is developed to explain how social media participate in tastemaking, namely social media's engines for 'making up taste.'

Problem Domain

The main problematic of this study was the overemphasis on the role of algorithms and user practices to explain how social media platforms altering cultural intermediation. I pointed

how in the current state of knowledge, the data processes preceding the work of algorithms are often neglected and/or taken for granted. I also problematized the ways in which user practices are assumed to be naturally occurring on social media, leaving a significant gap in our understanding of the implications of these sociotechnical arrangements. Departing from this point, I elaborated how both the user-practices and the work of recommendation algorithms are shaped by the distinct configuration of social media and discussed how this is implicated in the transformation of cultural intermediation. A shortage in studies recognising these aspects motivated this study. To address this gap I focused on the distinct relationship between the frontend and backend of social media and unpacked this relationship to explain how this relationship is implicated in decoding something as abstract as taste and turning it into data to produce personalised recommendations *matching* users' *tastes.*

Theory: Platformization

I presented a theoretical approach to address the identified gap. This approach has combined the concept of 'cultural intermediary' (Bourdieu, 1984) with the notion of 'platformization' (Constantinides, Henfridsson & Parker, 2018; Nieborg and Poell, 2018; van Dijck et al., 2018). This include conceptualising cultural intermediation as comprised of the processes mediating between the production and consumption of cultural content, key instances of which identified as curation, guidance and measurement. Platformization on the other hand conceptualised as the transformation of traditional value flows by digital platforms like social media. Drawing on these two theoretical approaches, the concept of 'platformization of cultural intermediation' is developed. This is the most significant contribution of this thesis. It advances the existing scholarship by explaining how social media internalize prevailing processes of cultural intermediation through the platform system in the course of their *development* and *optimisation*. It is demonstrated how in the course of being 'plat-formed' these processes *merge*, get *layered* and *shared* out by the users of social media and the platform itself.

Empirical Analysis and the Explanatory Framework

Drawing on the literature on platforms, the key aspects of *platformization* is identified as the *platform design* and *data operations*. Then its constituent elements are outlined to devise a preliminary conceptual framework to guide the analysis of the case study. In order to validate

the ideas put forward in devising the initial framework a social media platform is studied in-depth during the process of its development. To this end, a typical platform representative of social media operating in the domain of music discovery is selected for empirical investigation. The conceptual framework is applied to analyse the empirical data collected from multiple sources. Thus the platformization of music discovery is examined through a detailed analysis of the overall design logic and the underpinning data operations of the studied platform. The analysis demonstrated that curation, guidance, and measurement practices merge and performed by both the social media's users and its platform system at different layers. It is shown that this is enabled by the relationship between the *frontend* and *backend* of social media, thanks to the distinct logic underpinning its design and matchmaking operations. This is empirically illustrated by examining how a) users' interaction with the music is encoded on through the stylized actions organising user-platform participation at the frontend, b) how encoded data are then aggregated to construct digital version of listeners, 3) and finally how the platform computes listeners' music tastes to construct personalised recommendations assumed as 'matching' their tastes.

Moreover, findings revealed that platformization could not be explained by merely scrutinising the *development* of a social media platform. *Optimisation* of the platform found to play equally important role in this process. It is revealed that, this comprises a continuous process within which the data infrastructure underpinning the operations of a social media platform is continuously fine-tuned through data aggregation, normalisation, and synchronisation operations. Empirical findings demonstrated that this process renders a social media platform significantly reliant on the other platforms that operate within the same ecosystem, thereby operations of social media become partially contingent on other platforms. This highlighted how social media are not standalone entities but instead operating in a dynamic ecosystem in which various platforms operate interdependently in ways enabled by the infrastructural boundary making resources such as APIs and related data requirements. These findings validated previous views on how distinct platforms operating as 'microsystems' come together and constitute the 'ecosystem of connective media' (van Dijck, 2013), and how APIs are the significant enablers of this (Helmond, 2015; Plantin et al. 2016). What characterizes the ecosystem most, according to van Dijck (2013), is the interdependence and interoperability among these platforms. This thesis provided evidence confirming this by exposing the relationship among a set of internal and external sociotechnical elements underpinning the development and optimisation of a social media

platform. Findings highlighted how APIs play important roles in connecting different platforms within an ecosystem. However, understanding the dynamic of this ecosystem was beyond the scope of this thesis. Instead, this study focused on studying an individual platform in-depth. Given that the value of in-depth studies of microsystems is highly recognised (Van Dijck, 2013), this study made theoretical and empirical contributions to the literature through the empirical evidence collected from a microsystem. Nevertheless, empirical findings reveal that conducting research on the interdependence of platforms is needed to further advance our understanding of the social media phenomenon and its cultural implications.

Furthermore, optimisation process is observed to be remarkably reliant on data analytics tools operating in a cyclical fashion. Strategies embedded in the platform design are continuously fine-tuned to improve user engagement. It is shown that this is accomplished through strategically designed feedback loops that invite and exploit reactive tendencies of users to foster 'active' participation (Gerlitz and Lury, 2014). It is demonstrated that data analytics is cyclically used to spot user-engagement problems as well as to monitor how users react to the ways in which these problems are addressed by tweaking the platform design. Based on the analysis of this cyclical relationship, it is concluded that the platform and the user behaviour are mutually calibrated. The important point highlighted here was the ways in which social media platforms deliberately shape user behaviours in order to create 'a particular type of user' (van Dicjk, 2013). It is shown in detail how user behaviours are shaped in reaction to various technical and business related concerns. It is also emphasised that users are not just passive recipients of technological prescription. They find ways to resist to platform strategies (Bucher, 2012; Mager, 2012; Van Dicjk, 2013). Nevertheless the platform is tweaked again in response to user reactions thus shown how both the platform and user behaviours recursively shape each other. Scholars elaborated these aspects previously. This thesis empirically contributed to this literature by validating the existing views drawing on the empirical evidence collected from a social music platform. It also theoretically contributed to the literature by conceptualising the optimisation, and the use of data analytics tools in this process, as important aspects of how user behaviours are shaped to create desired outcomes. In this regard, the thesis advances the knowledge in this domain both theoretically and empirically.

Implications: Social Media as Sociotechnical Tastemakers

One of the most important contributions of this thesis is the conceptualization of social media as 'sociotechnical cultural intermediaries'. As discussed in the literature review, tastemaking is considered as a defining attribute of the traditional cultural intermediaries (Smith-Maguire and Matthew, 2014). From this perspective it is argued that as contemporary cultural intermediaries social media inherit this tastemaking quality. As *matchmakers* social media bring together the producers and consumers of cultural content in ways 'matching' their tastes. In doing so, they do not merely operate as matchmakers, but they also operate as *tastemakers* due to performative aspects of their data operations. This is justified in light of the previous studies that emphasized the self-fulfilling effects of quantification, categorization and measurement (Espeland and Stevens 2008, Callon 1998; MacKenzie 2006; Hacking 1986, 1995; 2002; 2007). These new kind of sociotechnical tastemakers acquire legitimacy from the longstanding trust in the objectivity of numbers (Desrosières 1998; Porter, 1995) and the processes of quantification, categorization, and measurement underpinning their production (Hacking 1986, 2007). This is in sharp contrast with the ways in which traditional tastemakers gain legitimacy through their tacit knowledge, expertise and other forms of social and cultural capital (Smith-Maguire and Matthew, 2014; Morris 2015a; Wright, 2015). In this regard, the data means social media use to measure user behaviours and define their taste through dynamically changing categories are conceptualised in this study as 'social media's engines' for 'making up taste'. This concept is the other theoretical contribution of the thesis. It is developed to emphasize how these data intensive processes are not free from consequences. This study provided evidence for this from the domain of culture. It is shown how users take up personalized recommendations (Beer 2013; Alaimo and Kallinikos 2017) and change their habitual behaviours either by following or resisting to the platform logic (Karakayali et al., 2017; van Dijck, 2013; Bucher 2012). As they do so, the ways in which knowledge produced about them through such data-means feeds back into their preferences, and eventually change them (Hacking, 2007; Karakayali et al., 2017). In other cases, knowledge produced about people through similar data means may also be used to discriminate them. As highlighted by scholars, these new forms of categorising and sorting out people determine how people are served, how they are treated, and/or who is looked upon favourably (Graham, 2004).

This thesis highlighted a minor instance of the possible implications of representing people through engineering devices and sorting them out by structuring their online behaviours. But even this minor instance can be used to discriminate people. This echoes, Crang and

Graham's (2007) views regarding how these organisations exercise power to discriminate: They "*track individuals to harvest data and make predictions to target the most appropriate or profitable consumer* " (p. 8011). This can be illustrated by how the music metadata platform Echo Nest, now acquired by Spotify, offered a number of criteria to the music services using its 'taste profile technology' in order to identify, acquire, engage and monetize what they refer to as 'high-value listeners'. The White Paper published by the company[11] describes this as the following: "*The majority of listening comes from a small segment of users: the committed, engaged, 'high-value' listeners that grow the business as opposed to the 'low-value' tourists who disappears from the sight".* These so called 'low-value' tourists are the 'non-paid users' who are described as "*inactive, thus contributing little advertising revenue.*" In this picture, successful services argued to be "*those that proactively predict which listeners are likely high-value, and focus on retaining and monetizing those listeners."* Arguably, The Echo Nest's taste profile technology *"analyses a service's users and identify the patterns of music taste that best predicts a user's future value"* by providing *" in-depth understanding of an individual listeners based on music listening behaviour, likes, dislikes, and other music activity."* It is argued that the company is able to identify psychographic/affinity characteristics of high-value listeners to help service monetize this group via targeted advertising. This indicates how taste profiles of users, constructed in the quite similar ways empirically demonstrated in this study, are explicitly used to discriminate listeners who are defined in particular systems as 'low-value' users.

It is remarkable to find that a music platform can publicly announce that it is able to "*identify music characteristics of high-value listeners to help service tailor user experience towards this high-value group*." This raises important questions regarding how these platforms are shaping cultural taste according to the commercial interests of these businesses: What is the consequences of being identified as a 'high-value' or a 'low-value' user by these services? What are the implications of the criteria these services use to represent their users in the form of data? How do they shape cultural experiences? How this is implicated in how people make decisions and how cultural tastes are shaped? These questions are far from being trivial given, in Beer's words, "*how these systems intervene in some of the social connections and*

[11] The Echonest (2013), How Music Services Can Acquire, Engage, and Monetize High-Value Listeners. Available at: http://static.echonest.com/WhitePapers/The%20Echo%20Nest%20-%20Monetizing%20High%20Value%20Listeners.pdf

divisions that we have long understood to be linked to the expressions of taste" (Beer, 2013: p. 96). This indicates that there is a significant transformation in the ways in which how culture is ordered today yet we are hardly aware of the implications: "*we need now to pay attention to what might be thought of as the infrastructures of taste formation and to 'a potential shift in the ontology of taste formation*" (ibid: p.97).

Cultural taste can be considered as a minor instance of how people may be discriminated considering how in certain occasions "being targeted and viewed preferentially might actually play closely with life opportunities, comfort, and social mobility, if, for example, the products are financial services, healthcare services, or insurance" (Beer, 2013). Nevertheless, another example from The Echo Nest indicates that constructing knowledge about cultural tastes may not be as neutral and innocent as we might imagine. In the past, the company used listening preferences to predict people's political affiliation[12] that they expressed through data such as their Facebook Likes. It is argued that, the company predicted that in the US, "Democrats liked more music genres (ten) than Republicans (seven); but liking The Beatles predicted nothing in the way of political preference." This indicates that we need to start asking critical questions regarding the consequence of these experiments with data generated through mundane and trivial daily activities, such as listening to music.

9.2. Contributions

As discussed in detail in the previous chapter, this thesis makes several empirical and theoretical contributions to advance the scholarship in the area of the research it addresses. First of all, the thesis makes a theoretical contribution by developing the concept of 'platformization of cultural intermediation' to study social media in cultural intermediary terms. An explanatory framework is developed how social media platforms internalize the curation, guidance and measurement processes that have long been under the control of professional authorities. This framework combined two theoretical lenses that focused on digital platforms and social media as their prominent instances. This framework outlined the *development* and *optimization* processes involved in building a digital platform. Parker et

[12] Wired (2014). "Echo Nest knows your music, your voting choice". Available https://www.wired.co.uk/article/echo-nest

al.'s (2016) general framework on the design logic of platforms and a more focused framework proposed by Alaimo and Kallinikos' (2017), outlining the data-based operations underpinning social media's matchmaking operations, are used complementarily. These frameworks are combined to overcome the limitation of Parker et al.'s (2016) framework regarding the ways in which data about users' taste are produced to be used in platforms' matchmaking operations. This initial framework delineated the key aspects of *development* of a social media platform. These ideas are refined in light of the empirical study and *optimisation* process of a social media platform identified as an equally important aspect of platformization.

This is one of the most important contributions of this thesis. This combination of two distinct theoretical approaches helped unpacking the configuration of social media platforms and revealed the sociotechnical elements involved in transformation of cultural intermediation. Considering how social media are prominent instances of platform-based business models, this thesis advances scholarship by contributing to the research on social and cultural implications of social media and digital platforms. Scholars explained cultural implications of social media by focusing on role of users and/or algorithms. However, as confirmed empirically, a range of sociotechnical factors preceding the work of algorithms and throughout the evolution of social media platforms collectively explain how these platforms shape user behaviours and their tastes in accordance with their business interests. In this regard, this combination of the two theoretical lenses used in this study can be employed in future studies in order to look beyond the overemphasized role of users and the work of algorithms and/or to explain implications of social media platforms in other domains.

Second theoretical contribution of thesis is conceptualising social media as 'sociotechnical cultural intermediaries'. It is discussed how these contemporary sociotechnical intermediaries owe their legitimacy as tastemakers from the assumed neutrality and objectivity of their data operations. In connection to this conceptualisation, the concept of 'social media's engines for making up taste' is developed to explain the performative implications of the often taken-for-granted matchmaking operations of social media. This concept advances scholarship by highlighting how the data-based means used to produce knowledge about people's cultural tastes recursively shape their taste in the course of describing it.

The thesis also made several empirical contributions advancing previous views. These are discussed in detail in the previous chapter. One noteworthy empirical contribution is providing evidence for how social media architect particular governance strategies (Tiwana 2013; Parker et al., 2016) to shape user behaviours. This study demonstrated how social media platforms intentionally use data-based means to keep their users engaged. They do so to generate 'active' user participation by inviting and exploiting the reactive tendencies of their users. This study empirically illustrated how social media platforms accomplish this using unique mechanism called the 'feedback loops' built on the strategies of *personalisation*, *metricisation*, and *notification*. From this vein, the thesis demonstrated how social media *shape behaviours*: a) by structuring actions through prescriptive design choices b) as users try to play with the outcomes of social media's data operations, shaped by this very design (Mager, 2012; Beer, 2013). These aspects indicate how commercial imperatives of social media, focused on producing value for the business as well their users result, lead to 'construction of a particular kind of user', one that is 'actively engaged' (Van Dijck 2013). Thus, this study empirically contributed to the debates concerning non-neutrality of social media by highlighting the role of *analytics* in 'creating' this kind of user in ways linked to how the platform system and the user behaviour are mutually calibrated with each other. This corroborates with Van Dijck's (2013, p.6) views. She argues, "*it is a common fallacy, to think of platforms as merely facilitating networking activities; instead, the construction of platforms and social practices is mutually constitutive*." This, according to Mackenzie (2006), indicates the 'performativity of code', that is, how ordering power of software is embedded in daily activities in ways that prescribe, shapes, and constitute behaviours in highly invisible ways (p.45).

9.3. Limitations and Recommendations for Future Research

In the previous section, I discussed the contributions of the thesis and discussed the implications of this study for future research. In this section I continue discussing how it could guide future studies by pointing out the areas that require further research by reflecting on the limitations of the thesis. The first limitation concerns the chosen methodology. This research is built on a single explanatory case study (Yin, 2003). A single case study provided the researcher with the opportunity to have an in-depth understanding of the issues and concerns involved in the development of a social media platform in its real-life context. This

allowed advancing scholarship by providing detailed empirical evidence for other scholars' conceptual propositions and making new theoretical contributions. Empirical data underpinning these contributions were collected from an individual social media platform operating in the domain of music. However, the concepts developed in this study can be used in more generic terms in future research to study platformization of other domains of everyday life. To suggest a few potential avenues for further enquiry, one possible direction for future research would be the application of the framework on findings derived from another in-depth case study on a social media platform facilitating the exchange of other forms of cultural content. This can help improving knowledge in this domain by validating or falsifying the findings of this research. Thus alternative explanations can be developed. This may be methodologically complemented by conducting qualitative research with the users of this platform to assess how personalised recommendations shape their cultural experiences and involved in the formation of their tastes. This may include studying to what extent users value these recommendations and consider them credible. A noteworthy point to assess would be exploring how reflexive and aware are social media users about the data they provide about their cultural consumption and how this is involved in their experiences.

An alternative route might be conducting multiple case studies and comparing findings with each other. This helps to overcome the limitation arising from conducting a single case study. A comparative case study can focus on testing the explanatory power of the conceptual framework proposed in this study by comparing several social media specialized in music or other cultural domains. Furthermore, the scope of the research can be extended beyond social media. Other kinds of digital platforms (i.e. e-commerce or streaming platforms) can be examined by slightly modifying the same framework to systematize the study of platformization. This would help to unpack how digital platforms other than social media internalise curation, guidance and measurement given they are not reliant on users for these operations. The same study may also include the analysis of the governance strategies embedded in the design of these platforms to see whether they are using the same mechanisms to incentive and control user-platform participation.

We can also study platformization of other kinds of linear value flows of the past and examine how these processes are internalised by digital platforms (i.e. platformization of cab-hailing, holiday booking etc.). This is of particular relevance given the growing penetration of digital platforms in different avenues of daily life and result in a web of interconnected

platforms infrastructuring everyday living (Alaimo and Kallinikos, 2017; Constantinides, Henfridsson and Parker, 2018; Helmond, 2015; Plantin et al. 2016). Considering that the majority of the existing research on platforms and platformization is on established platforms, there is a recognized need for further empirical research unpacking *how exactly* platforms become platforms (Islind, 2018). In this regard, the theoretical approach presented in this study may guide future research by highlighting the key aspects of the development and the optimization of platforms other than social media and the ways in which they plat-form traditional value flows. Value of in-depth studies of individual platforms is recognized in the literature. However, scholars emphasise that there is a growing need to expose how different platforms constitute an ecosystem and operate interdependently (van Dicjk, 2013). This study empirically showed how different platforms operating in the same ecosystem rely on each other's data. Future studies may advance scholarship by focusing these aspects to empirically demonstrate how different platforms shape each other and evolve together in a dynamic ecosystem.

Another limitation concerns the size of the studied company. Empirical findings were derived from a company in its start-up phase. Therefore, data was mainly collected through the interviews done with five people involved in the development of the platform chosen for empirical inquiry. Researching a start-up had several benefits and drawbacks. The main advantage was the ability to collect data directly from the developers who were responsible from all the operations involved in running a company. Given that they are the expert developers of the object of the study, the researcher had the chance to construct the case study narrative drawing on the first-hand experience of the experts of the studied topic. This included gathering information about the platform's development phase in light of both the technical and the business concerns. Deriving such a narrative in a more established company with a more hierarchical organizational structure would have been more challenging. In this regard, future studies may turn such limitations into advantage unpack how exactly platforms become platforms and how they realise platformization.

The final point that is noteworthy to emphasize is the advantages and disadvantages of researching a start-up company. Start-ups operate in highly agile manners. The flexibility of running a start-up company gives the co-founders the ability to respond the external changes very quickly. That is they can rapidly adjust to changes according to the perceived user needs as well as the requirements of the ecosystem within which they operate. They could thus alter

the platform configuration several times within very short time spans. This creates several challenges and opportunities for research. The most prominent challenge for this study was to the ambiguities involved in studying such a dynamically changing artifact. This necessitated the researcher to adjust the corpus design three times in accordance with the changes in platform configuration. These changes also paved the way for further development of the explanatory framework in light of the unexpected findings derived from the empirical study. This indicates how agility of startups provides researchers the opportunity to observe complex sociotechnical phenomena in real-life context as they unfold in relatively short periods of time. Making similar observations within the context of a more established platform business may be way more difficult. In this regard, this limitation related to the size of the studied company can also be considered as an opportunity, which allows revealing certain aspects of platforms and platformization that would not be possible to show in other settings.

Appendix

A: Interviews

Name of the Interviewee	Interview Dates	Role
Onur Y.	5 April 2014 8 May 2014 18 July 2014 14 November 2014	CEO/Cofounder/Developer
Ozgur	19 April 2014 14 August 2014 13 November 2014 18 December 2015	Cofounder/Developer
Baris A.	7 April 2014 18 April 2014 18 July 2014 18 December 2015	Cofounder/Business Analyst (User Experience, PR and Marketing)
Erdem G.	29 April 2014 5 June 2014 27 June 2014 11 July 2014 18 July 2014 14 August 2014 26 March 2015 (via Skype) 16 December 2015	Cofounder/Business Analyst (User Experience, PR and Marketing)
Rifat O.	27 June 2014	Mobile App Developer
Joint Interview with Baris A. and Erdem Gelal	14 May 2014 27 June 2014 14 August 2014	Analysts
Team Meeting	7 April 2014 26 December 2014	

B: List of Video Content

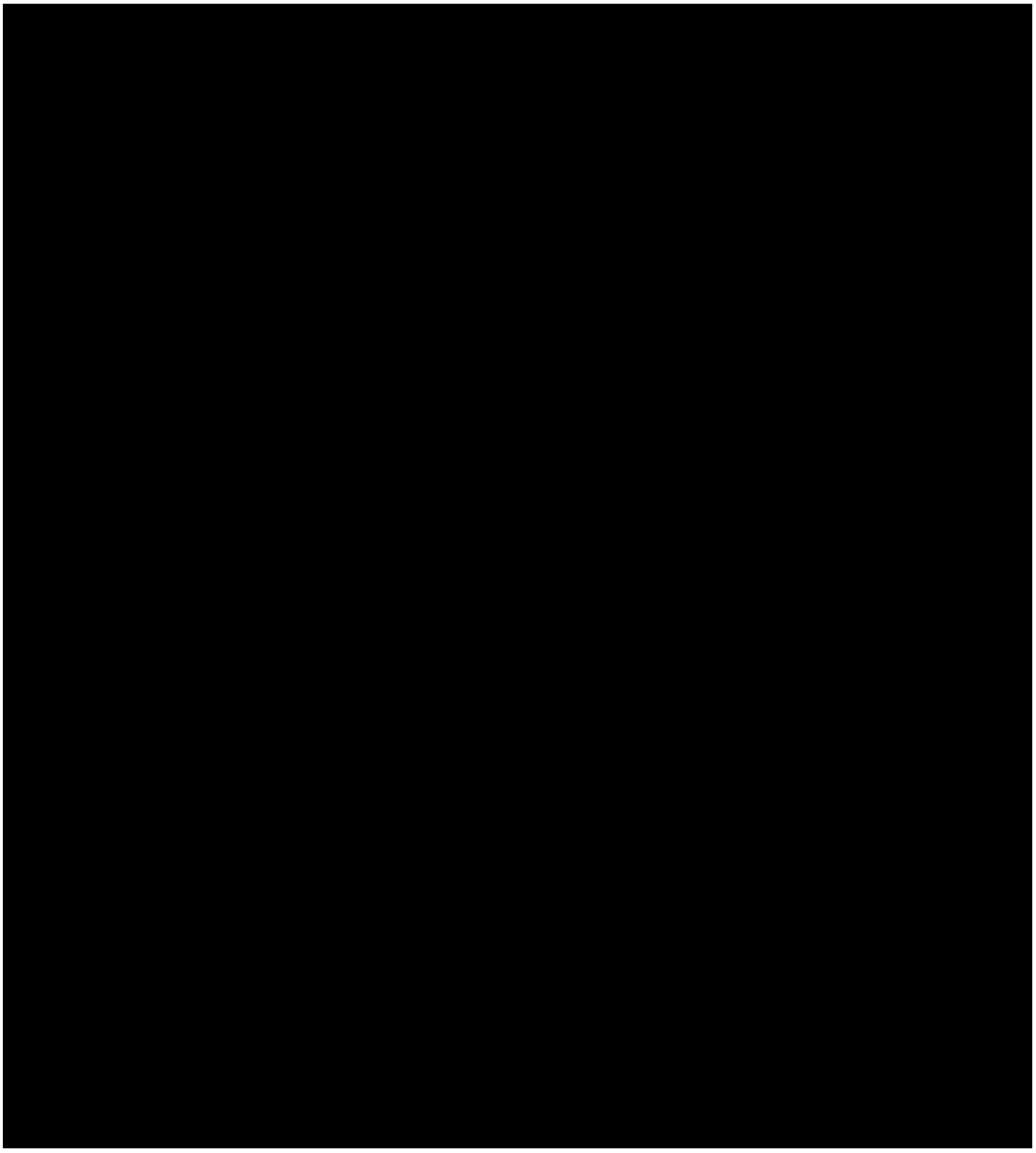
ID	Name / Description	Link	Data Published
1	Cubic.fm Product Video	https://www.youtube.com/watch?v=O1nV8pS2Ke0	April 16, 2014
2	Cubic.fm Demo Video	https://www.youtube.com/watch?v=qd7cGUSKMRM&t=1s	June 3, 2015
3	Cubic.fm, Startup Istanbul 2014	https://www.youtube.com/watch?v=VjlR-NS9aPI	April 28, 2015
4	2015 Etohum 15 Girişimleri: Cubic.FM (Entrepreneur Interviews: Cubic.fm)	https://www.youtube.com/watch?v=okALCakQgjw	Feb 11, 2015
5	Cubic.fm'den Erdem Gelal, Startup Turkey 2015 Erdem Gelal from Cubic.fm interview/ Startup Turkey	https://www.youtube.com/watch?v=LEvJQGw27wM	May 26, 2015
6	Startup Turkey Challenge 2015 Üçüncüsü: Cubic.fm (Cubic.fm 3rd winner of Startup Turkey Challenge 2015)	https://www.youtube.com/watch?v=BpOuk9dHI04	March 3, 2015
7	Türk Telekom - PİLOTT | Girişimci Röportajları | Barış Can Aktepe | Cubic.fm (Entrepreneur interviews –Baris Can Akcatepe from Cubic.fm)	https://www.youtube.com/watch?v=LhDMeUH-o0c	Feb 4, 2014

8	#32 cubic.fm ekibinden erdem gelal ile müzik ve teknoloji konuşuyoruz	https://www.youtube.com/watch?v=Zb59z1GWGbM&t=1s	October 15, 2015
9	TNW - Ayelet Noff - Pitch Competition | The Next Web	https://[redacted].youtube.com/watch?v=om_YrZshMls&t=4644s	April 24,2014
10	Midemlab: Music Discovery, Recommendation & Creation pitches - Midem 2014	https://www.youtube.com/watch?v=Xx_90ag8Oqk	Feb 1, 2014
11	Cubic.fm	https://vimeo.com/96534333	26 May, 2014

C: Examples of Documents

A Sample SQL Query of Songs kept in the database with data points aggregated from EchoNest and user-generated tags aggregated from the Last.fm

A News Article about the company



A blog post about the company

A diagram sketched by the researcher to visualise construction of personalised recommendations for crosschecking with the developers.

A screenshot from the launch of the private beta in ProductHunt website for feedback and Q&A from the developer community and angel investors. Available at: https://www.producthunt.com/posts/cubic-fm-2

Sample Screenshot from the Platform's Support Portal for Users

Sample Screenshot from the Platform's Support Portal for Users

Sample Screenshot from the Platform's Support Portal for Users

Sample Screenshot from the Platform's Support Portal for Users

Sample Screenshot from the Platform's Support Portal for Users

References

Aaltonen, A., 2011. *Manufacturing the digital advertising audience* (Doctoral dissertation, The London School of Economics and Political Science (LSE)).

Aaltonen, A 2015, 'Making Up Digital Advertising Audiences: The New Form of Contracting Human Attention', *paper presented at 31st European Group for Organizational Studies (EGOS) Colloquium*, Athens, Greece, July 2–4, 2015

Aaltonen, A. and Tempini, N., 2014. Everything counts in large amounts: a critical realist case study on data-based production. *Journal of Information technology*, *29*(1), pp.97-110.

Alaimo, C., 2014. *Computational Consumption: Social media and the construction of digital consumers* (Doctoral dissertation, The London School of Economics and Political Science (LSE).

Alaimo, C. and Kallinikos, J., 2016. Encoding the everyday: The infrastructural apparatus of social data. *Big Data Is Not a Monolith*, p.77.

Alaimo, C. and Kallinikos, J., 2017. Computing the everyday: Social media as data platforms. *The Information Society*, *33*(4), pp.175-191.

Alasuutari, P., Bickman, L., & Brannen, J. (Eds.). (2008). *The SAGE handbook of social research methods*. Sage.

Alvesson, M. and Sköldberg, K., 2009. *Reflexive methodology: New vistas for qualitative research*. Sage.

Ang, I. (1991). *Desperately seeking the audience*. Routledge.

Apperley, T., & Parikka, J. (2018). Platform studies' epistemic threshold. *Games and Culture*, *13*(4), 349-369.

Aral, S., Dellarocas, C. and Godes, D., 2013. Introduction to the special issue—socialmedia and business transformation: a framework for research.*Information Systems Research*, *24*(1), pp.3-13.

Arsel, Z. and Bean, J., 2012. Taste regimes and market-mediated practice. *Journal of Consumer Research*, *39*(5), pp.899-917.

Ashton D and Couzins M (2015) Content curators as cultural intermediaries: 'My reputation as a curator is based on what I curate, right?' M/C – A Journal of Media and Culture 18(4).

Baldwin, C.Y., Woodard, J.J., 2009. The architecture of platforms: a unified view. In: Gawer, A. (Ed.), Platforms, Markets and Innovation. Edward Elgar, Cheltenham, UK and Northampton, Mass, pp. 19–44.

Baldwin, C.Y., Clark, K.B., 2000. Design Rules: The Power of Modularity, vol. 1. MIT Press, Cambridge, MA.

Balzer D (2014) Curationism: How Curating Took Over the Art World and Everything Else. Toronto: Couch House Books.

Barna, E. (2017). "The perfect guide in a crowded musical landscape:" Online music platforms and curatorship. *First Monday*, *22*(4).

Bauer, M. W., and Aarts, B. 2000. Corpus construction: a principle for qualitative data collection. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative researching with text, image and sound. A practical handbook for social researchers* (pp. 19-37). London: Sage.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report, 13* (4), 544-559.

Baym, N. K. (2010). Personal Connections in the Digital Age: Wiley.

Baym, N. K. (2013). Data not seen: The uses and shortcomings of social media metrics. *First Monday, 18*(10).

Beer, D. D. (2008). Social network (ing) sites… revisiting the story so far: A response to danah

boyd & Nicole Ellison. *Journal of Computer-Mediated Communication, 13*(2), 516-529.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society, 11*(6), 985-1002.

Beer, D. (2013). *Popular culture and new media: The politics of circulation*. Springer.

Beer, D. (2015). Productive measures: Culture and measurement in the context of everyday neoliberalism. *Big Data & Society*, *2*(1), p.2053951715578951.

Beer, D., and Burrows, R. (2010). Consumption, Prosumption and Participatory Web Cultures: An Introduction. Journal of Consumer Culture 10: 3–12

Beer, D., and Burrows, R. (2013). Popular Culture, Digital Archives and the New Social Life of Data. Theory, Culture & Society, 1-25.

Beer, D. and Taylor, M. (2013). The hidden dimensions of the musical field and the potential of the new social data. *Sociological Research Online*, *18*(2), p.14.

Benghozi, P. J., & Paris, T. (2016). The cultural economy in the digital age: A revolution in intermediation?. *City, Culture and Society, 7*(2), 75-80.

Berger, P.L. and Luckmann, T., 1991. *The social construction of reality: A treatise in the sociology of knowledge* (No. 10). Penguin UK.

Bermejo, F., 2007. *The internet audience: Constitution & measurement*. Peter Lang.

Bermejo, F., 2009. Audience manufacture in historical perspective: from broadcasting to Google. *New Media & Society*, *11*(1-2), pp.133-154.

Bhaskar, M., 2016a. *Curation: The power of selection in a world of excess*. Hachette UK.

Bhaskar, M., 2016b.' In the age of the algorithm, the human gatekeeper is back.' *The Guardian,* 30 September

Biber, D., 1993. Representativeness in corpus design. *Literary and linguistic computing*, *8*(4), pp.243-257.

Bista, S.K., Nepal, S., Colineau, N. and Paris, C., 2012, October. Using gamification in an online community. In *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2012 8th International Conference on* (pp. 611-618). IEEE.

Bogdan, R. C., & Biklen, S. K. (1992). Qualitative research: An introduction to theory and methods. *Needham Height: Allyn & Bacon*.

Bogner, A. and Menz,W. (2002) "Das theoriegenerierende Experteninterview - Erkenntnisinteresse, Wissensform, Interaktion," in A. Bogner, B. Littig, and W. Menz (eds.), Das Experteninterview - Theorie, Methode, Anwendung. Opladen: Leske & Budrich. pp. 33—70.

Bogost, I., & Montfort, N. (2007). New media as material constraint. An introduction to platform studies. In *Electronic Techtonics: Thinking at the Interface. Proceedings of the First International HASTAC Conference* (pp. 176-193).

Bogost, I., & Montfort, N. (2009). Platform studies: Frequently questioned answers.

Boudreau, K. J., & Hagiu, A. (2009). Platform rules: Multi-sided platforms as regulators. *Platforms, markets and innovation*, *1*, 163-191.

Bourdieu, P., 1984. *Distinction: A social critique of the judgement of taste*. Harvard university press.

Bourdieu, P. (1996). *Photography: A Middle-Brow Art*. Cambridge: Polity Press.

Bowker, G.C. and Star, S.L., 1999. *Sorting things out: Classification and its consequences*. MIT press.

Bowker, G.C. and Star, S.L., 2000. Invisible mediators of action: Classification and the ubiquity of standards. *Mind, Culture, and Activity*, *7*(1-2), pp.147-163.

Boyatzis, R.E., 1998. *Transforming qualitative information: Thematic analysis and code development*. sage.

boyd, d. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship.*Journal of Computer-Mediated Communication, 13* 210-230.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, *15*(5), 662-679.

Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), pp.77-101.

Bruns, A.(2011). Gatekeeping, gatewatching, real-time feedback: New challenges for journalism. Brazilian Journalism Research 7(10): 117–136.

Bruns, A. (2015). Working the story: News curation in social media as a second wave of citizen journalism. *The Routledge companion to alternative and community media*, 379-388.

Bruns, A. (2018). *Gatewatching and news curation: Journalism, social media, and the public sphere*. Peter Lang.

Bucher, T., 2012. Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *new media & society*, *14*(7), pp.1164-1180.

Bucher, T., & Helmond, A. (2017). The affordances of social media platforms. *The SAGE handbook of social media*, 223-253.

Bunchball, I., 2010. Gamification 101: An introduction to the use of game dynamics to influence behavior. *White paper*, p.9.

Burgess, J., Baym, N., Bucher, T., Helmond, A., John, N. A., Nissenbaum, A., & Craig, D. (2017). Platform studies: The rules of engagement.

Burgess, J., Marwick, A., & Poell, T. (Eds.). (2017). *The SAGE Handbook of Social Media*. Sage.

Butler, J (1997) Excitable Speech: A Politics of the Performative, London: Routledge.

Çalışkan, K., & Callon, M. (2010). Economization, part 2: a research programme for the study of markets. *Economy and Society*, *39*(1), 1-32.

Callon, M. (1998). The Laws of the Markets. Oxford: Blackwell

Callon, M. (2007). What does it mean to say that economics is performative? In MacKenzie, D., Muniesa, F and Siu, L. (eds), Do Economists Make Markets? On the Performativity of Economics. Princeton, NJ: Princeton University Press.

Callon, M., Millo, Y., & Muniesa, F. (2007). *Market devices*(No. halshs-00177891).

Chen, Z. T. (2018). Poetic prosumption of animation, comic, game and novel in a post-socialist China: A case of a popular video-sharing social media Bilibili as heterotopia. *Journal of Consumer Culture*, 1469540518787574.

Cheney-Lippold, J. (2011) 'A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control', Theory, Culture & Society 28(6): 164–181.

Clark, J., Couldry, N., De Kosnik, A.T., Gillespie, T., Jenkins, H., Kelty, C., Papacharissi, Z., Powell, A. and van Dijck, J., 2014. Participations| Part 5: PLATFORMS. *International Journal of Communication*, 8, 28

Constantinides, P., Henfridsson, O. and Parker, G.G. (2018) Introduction—Platforms and Infrastructures in the Digital Age. Information Systems Research 29(2):381- 400.

Corbin, J. and Strauss, A. (2008). Basics of qualitative research: Techniques and procedures for developing grounded theory.

Couldry N and Kallinikos J (2017) Ontology. In: Burgess J, Marwick A and Poell T (eds) The SAGE Handbook of Social Media. London: Sage, pp.146–159.

Crang, M. & Graham, S. (2007) 'Sentient Cities: Ambient Intelligence and the Politics of Urban Space', Information, Communication & Society 10(6): 789–817.

Crotty, M., 1998. *The foundations of social research: Meaning and perspective in the research process*. Sage.

Dalton, C., & Thatcher, J. (2014). What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'big data'. *Society and Space, 29*.

Denzin, N.K. (1989b) The Research Act (3rd edn). Englewood Cliffi, NJ: Prentice Hall.

de Reuver, M., Sørensen, C. and Basole, R.C., 2016. The digital platform: a research agenda. *Journal of Information Technology*, pp.1-12.

Desrosières A (1998) The Politics of Large Numbers: A History of Statistical Reasoning, trans. Camille Naish. Cambridge, MA/London: Harvard University Press.

Deterding, S., Dixon, D., Khaled, R. and Nacke, L., 2011, September. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15). ACM.

Dörr, J., Wagner, T., Benlian, A. and Hess, T., 2013. Music as a Service as an Alternative to Music Piracy?. *Business & Information Systems Engineering*,*5*(6), pp.383-396.

Drabek, M.L., 2009. Interactive classification and practice in the social sciences: expanding Ian Hacking's treatment of interactive kinds. *Issues in the Rhetoric of Scienceand Technology*, 6(2), pp. 62-80.

Eisenmann, T., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two-sided markets. *Harvard business review*, *84*(10), 92.

Elias, N., 1994. The civilizing process, trans. Edmund Jephcott. *Oxford: Blackwell*, *65*, p.1.

Ellison, N. B., & Vitak, J. (2015). Social network site affordances and their relationship to social capital processes. *The handbook of the psychology of communication technology*, 205-227.

Espeland, W.N. and Sauder, M., 2007. Rankings and reactivity: How public measures recreate social worlds. *American journal of sociology*, *113*(1), pp.1-40.

Espeland, W.N. and Stevens, M.L., 2008. A sociology of quantification.*European Journal of Sociology/Archives Européennes de Sociologie*, *49*(3), pp.401-436.

Ettema, J.S. and Whitney, D.C., 1994. *Audiencemaking: How the media create the audience* (Vol. 22). Sage Publications, Inc.

Evans, D.S., 2003. Some empirical aspects of multi-sided platform industries. *Review of Network Economics*, *2*(3).

Evans, D. S., & Schmalensee, R. (2007). *Catalyst code: the strategies behind the world's most dynamic companies*. Harvard Business School Press.

Evans, D.S. and Schmalensee, R., 2016. *Matchmakers: the new economics of multisided platforms*. Harvard Business Review Press.

Featherstone, M. (1991). *Consumer Culture and Postmodernism*. London: Sage.

Fergie, G., Hunt, K., & Hilton, S. (2016). Social media as a space for support: young adults' perspectives on producing and consuming user-generated content about diabetes and mental health. *Social Science & Medicine*, *170*, 46-54.

Flick, U. (2004). Triangulation in qualitative research. *A companion to qualitative research*, *3*, 178-183.

Flick, U. (2009). *An introduction to qualitative research* (4th ed.). London: Sage.

Floridi, L. (Ed.). (2008). *The Blackwell guide to the philosophy of computing and information*. John Wiley & Sons.

Floridi, L. (2010). *Information: A very short introduction*. OUP Oxford.

Foster, P., Borgatti, S.P. and Jones, C. (2011). Gatekeeper search and selection strategies: Relational and network governance in a cultural market. Poetics 39(4): 247–265.

Foucault, M. (1970) The Order of Things: An Archaeology of Human Sciences. London:Tavistock.

Foucault M (2007) Security, Territory, Population, ed. Arnold Davidson, trans. Graham Burchell. Lectures at the Collège De France, 1977–78. Basingstoke: Palgrave Macmillan.

Fuchs, C. (2012). Dallas Smythe Today-The Audience Commodity, the Digital Labour Debate, Marxist Political Economy and Critical Theory. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, *10*(2), 692-740.

Fuchs, C. (2014). Digital prosumption labour on social media in the context of the capitalist regime of time. *Time & Society*, *23*(1), 97-123.

Gaskell, G., 2000. Individual and Group Interviewing. In M. W. Bauer & G. Gaskell, eds. *Qualitative researching with text, image and sound*. London et al.: Sage, pp. 38–56.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research policy*, *43*(7), 1239-1249.

Gerlitz, C. and Helmond, A., 2013. The like economy: Social buttons and the data-intensive web. *New Media & Society*, *15*(8), pp.1348-1365.

Gerlitz, C. and Lury, C., 2014. Social media and self-evaluating assemblages: On numbers, orderings and values. *Distinktion: Scandinavian Journal of Social Theory*, *15*(2), pp.174-188.

Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third- party development: the boundary resources model. *Information Systems Journal*, *23*(2), 173-192.

Gillespie, T., 2010. The politics of 'platforms'. *New Media & Society*, *12*(3), pp.347- 364.

Gillespie, T., 2014. The relevance of algorithms. *Media technologies: Essays on communication, materiality, and society*, *167*.

Gillespie, T., 2018. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Gitelman, L. ed., 2013. *Raw data is an oxymoron*. MIT Press.

Gitelman, L., & Jackson, V. (2013). Introduction. In "Raw data" is an oxymoron, ed. Lisa Gitelman, 1–14.

Graham, S. (2004) 'The Software-Sorted City: Rethinking the "Digital Divide" ', in Graham, S. (ed.), The Cybercities Reader, London: Routledge, pp. 324–332

Grosser, B., 2014. What do metrics want? How quantification prescribes social interaction on Facebook. *Computational Culture: a journal of software studies*, *4*.

Gubrium, J. F., & Holstein, J. A. (2001). *Handbook of interview research: Context and method*. Sage Publications.

Gummesson, E. (1991) Qualitative Methods in Management Research. Sage Publications. Newbury Park, California.

Hacking,I., 1986. Making up people. In T. C. Heller,M. Sosna,D.E. Wellbery,A.I. Davidson, A. Swidler, & I. Watt (Eds.), Reconstructing individualism: Autonomy, individuality, and the self in Westernthought(pp. 222-236). Stanford, CA: Stanford University Press,

Hacking, I., 1995. The looping effects of human kinds. *Causal cognition: A multidisciplinary debate*, *12*, pp.351-394.

Hacking, I., 1999. *The social construction of what?*. Harvard university press.

Hacking, I., 2002. Historical Ontology. Cambridge MA: Harvard University Press.

Hacking, I., 2007. *'Kinds of People: Moving Targets'*, Proceedings of the British Academy, 151, 285-318.

Hagiu, A., & Eisenmann, T. (2007). A staged solution to the catch-22. *Harvard Business Review*, *85*(11), 25-26.

Hamari, J. and Koivisto, J. (2013). Social Motivations To Use Gamification: An Empirical Study Of Gamifying Exercise. In *ECIS* (p. 105).

Hammersley, M. (2010). Reproducing or constructing? Some questions about transcription in social research. *Qualitative research*, *10*(5), 553-569.

Hartley, Jean (1994). Case studies in organizational research. In Catherine Cassell & Gillian Symon (Eds.), Qualitative methods in organizational research, a practical guide (pp.208-229). London: Sage

Helmond, A., 2015. The platformization of the web: Making web data platform ready. *Social Media+ Society*, *1*(2), p.2056305115603080.

Hendricks JM (2015) Curating value in changing markets: Independent record stores and the vinyl record revival. Sociological Perspectives 59(2): 479–497

Henfridsson, O. and Bygstad, B., 2013. The generative mechanisms of digital infrastructure evolution. *MIS quarterly*, *37*(3).

Hirsch, P.M., 1972. Processing fads and fashions: An organization-set analysis of cultural industry systems. *American journal of sociology*, *77*(4), pp.639-659.

Holstein, J.A. & Gubrium, J.F., 1997. Active Interviewing. In D. Silverman, ed. *Qualitative Research: Theory, Method and Practice*. London: Sage, pp. 113– 129

Hracs, B. J. (2015). Cultural intermediaries in the digital age: The case of independent musicians and managers in Toronto. Regional Studies 49: 461–475.

Hutchinson, J., 2015. "From Fringe to Formalisation: An Experiment in Fostering Interactive Public Service Media." Media International Australia 155: 5–15.

Islind, A. S. (2018). *Platformization: Co-Designing Digital Platforms in Practice* (Doctoral dissertation, University West).

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*.

Jansson, J. and Hracs, B. J. (2017). Curating Music in the Digital Age. *Economy, Governance, Culture 1/2017*.

Jansson, J. and Hracs, B. J. (2018). Conceptualizing Curation in the Age of Abundance: The case of recorded music. *Environment and Planning A: Economy and Space*, 0308518X18777497.

Johnson, R. B. (1997). Examining the validity structure of qualitative research. *Education*, *118*(2), 282.

Jenkins, H., Ford, S. and Green, J., 2013. *Spreadable media: Creating value and meaning in a networked culture*. NYU press.

Kallinikos, J., 2004. *Farewell to constructivism: Technology and context-embedded action* (pp. 140-162). Oxford University Press.

Kallinikos, J., 2006. *The consequences of information: institutional implications of technological change*. Cheltenham: Edward Elgar.

Kallinikos, J., 2009. On the computational rendition of reality: Artefacts and human agency. *Organization*, *16*(2), pp.183-202.

Kallinikos, J., 2013. The allure of big data. *Mercury Magazine*, *2*(3), pp.40-43.

Kallinikos, J., Aaltonen, A. and Marton, A., 2013. The Ambivalent Ontology of Digital Artifacts. *Mis Quarterly*, *37*(2).

Kallinikos, J. and Constantiou, I.D., 2015. Big data revisited: a rejoinder. *Journal of Information Technology*, *30*(1), pp.70-74.

Kallinikos, J. and Tempini, N., 2011. Post--material Meditations: On Data Tokens, Knowledge and Behaviour. In *27th EGOS Colloquium--European Group of Organizational Studies, Gothenburg, Sweden, July*.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 53, 59-68.

Karakayali, N., Kostem, B., & Galip, I. (2018). Recommendation Systems as Technologies of the Self: Algorithmic Control and the Formation of Music Taste. *Theory, Culture & Society*, *35*(2), 3-24.

Kassabian, A., 2001. Ubiquitous listening and networked subjectivity. *Echo: A Music–Centered Journal*, *3*(2).

Kassabian, A., 2013. *Ubiquitous listening: affect, attention, and distributed subjectivity*. Univ of California Press.

Katz, M. L., & Shapiro, C. (1986). Technology adoption in the presence of network externalities. *Journal of political economy*, *94*(4), 822-841.

Kelly, A., 2014. “Advertising.” In The Cultural Intermediaries Reader, edited by Jennifer Smith Maguire and Julian Matthews, 67–76. London: Sage

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons*, *54*(3), 241-251.

Kitchin, R. (2014a). Big Data, new epistemologies and paradigm shifts. Big Data & Society, 1(1), 2053951714528481.

Kitchin, R. (2014b). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

Kitchen, R., & Dodge, M. (2011). Code/Space. *Software and Everyday Life. Cambridge*.

Knafl, K. A., & Breitmayer, B. J. (1989). Triangulation in qualitative research: issues of conceptual clarity and purpose. In (Ed. JM Morse). Qualitative Nursing Research asContemporary Dialogue. Rockville: MD.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American journal of occupational therapy*, *45*(3), 214-222.

Kvale, S. & Brinkmann, S., 2008. *InterViews*, Thousand Oaks: Sage.

Lamere, P. (2008). Social tagging and music information retrieval. *Journal of new music research*, *37*(2), 101-114.

Lampe, C., 2014. Gamification and social media. *The Gameful world: Approaches, issues, applications. Cambridge, MA: MIT*.

Lange, B. (2016). The evolution of music tastemakers in the digital age: The rise of algorithms and the response of journalists. In *The Production and Consumption of Music in the Digital Age* (pp. 251-261). Routledge.

Langley, P., & Leyshon, A. (2017). Platform capitalism: the intermediation and capitalisation of digital economic circulation. *Finance and society.*, *3*(1), 11-31.

Langlois, G. and Elmer, G., 2013. The research politics of social media platforms. *Culture Machine*, *14*.

Lash, S. (2007) ‘Power after hegemony: Cultural studies in mutation’, Theory, Culture & Society 24(3): 55–78.

Lauriault, T.P. (2012) Data, Infrastructures and Geographical Imaginations. Ph.D. Thesis, Carleton University, Ottawa.

Law, J., 2004. *After method: Mess in social science research*. Routledge.

Lee, A.S. and Baskerville, R.L., 2003. Generalizing generalizability in information systems research. *Information systems research*, *14*(3), pp.221-243.

Lena, J.C., 2012. *Banding together: How communities create genres in popular music*. Princeton University Press.

Leslie, D., Brydges, T. and Brail, S. (2015). Qualifying aesthetic values in the experience economy: The role of independent fashion boutiques in curating slow fashion. In: Lorentzen A, Schrøder L and Topsø LK (eds) Spatial Dynamics in the Experience Economy. New York: Routledge, pp.88–102.

Levenshus, A. (2010). Online relationship management in a presidential campaign: A case study of the Obama campaign's management of its internet-integrated grassroots effort. *Journal of Public Relations Research*, *22*(3), 313-335.

Lewis, T., 2014. "Lifestyle Media." In The Cultural Intermediaries Reader, edited by Jennifer Smith Maguire and Julian Matthews, 134–144. London: Sage.

Liu, H., Maes, P. and Davenport, G., 2008. Unraveling the taste fabric of social networks. *International Journal on Semantic Web and Information Systems*, *2*(1).

Lovink, G. and Rasch, M., 2013. *Unlike us reader: Social media monopolies and their alternatives* (No. 8). Institute of Network Cultures.

Lu, X., Goldsmith, R.E. and Pagani, M., 2013. *Two-sided markets and social media* (pp. 197-213). Hershey, PA: IGI Global Press.

MacKenzie, D. (2006). An Engine, Not a Camera. Cambridgr, MA: MIT Press.

Mager, A. (2012). Algorithmic ideology: How capitalist society shapes search engines. *Information, Communication & Society*, *15*(5), 769-787.

Marshall, C., & Rossman, G. B. (1995). Data collection methods. *Designing qualitative research*, *2*(8).

Merton, R. K., & Kendall, P. L. (1946). The focused interview. *American journal of Sociology*, *51*(6), 541-557.

Meuser, M. and Nagel, U. (2002) "Expertlnneninterviews - vielfach erprobt, wenig bedacht. Ein Beitrag zur qualitativen Methodendiskussion," in A. Bogner, B. Littig, andW Menz (eds.), Das Experteninterview. Opladen: Leske & Budxich. pp. 71-95.

Mayer-Schönberger, V., & Cukier, K. (2013). Big Data: A revolution that transforms how we work, live, and think.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard business review*, *90*(10), 60-68.

Meehan, E.R., 1984. 'Ratings and the Institutional Approach: A Third Answer to the Commodity Question', *Critical Studies in Mass Communication* 1(2): 216–25.

Merriam, S. B., 2009. Qualitative research: A guide to design and implementation. San Francisco, CA: Jossey-Bass.

Miles, M.B. and Huberman, A.M., 1994. *Qualitative data analysis: An expanded sourcebook*. sage.

Moor, L., 2012. Beyond cultural intermediaries? A socio-technical perspective on the market for social interventions. *European Journal of Cultural Studies*,*15*(5), pp.563-580.

Morris, J., 2011. Sounds in the cloud: Cloud computing and the digital music commodity. *First Monday, 16*(5).

Morris, J.W., 2015a. Curation by code: Infomediaries and the data mining of taste. *European Journal of Cultural Studies*, *18*(4-5), pp.446-463.

Morris, J.W., 2015b. *Selling Digital Music, Formatting Culture*. Univ of California Press.

Mueller, K. (2012). Curators and Gatekeepers: Who Do You Trust? [Blog] *Business 2 Community.* Available at: http://www.business2community.com/social- media/curators-and-gatekeepers-who-do-you-trust-0236951#qp3ZYAp5AhzFKDEd.97 [Accessed 6 June 2016].

Müller, K.H., 2008. Methodologizing Radical Constructivism. Recipes for RC-Designs in the Social Sciences.

Muniesa, F., Millo, Y., & Callon, M., 2007. An introduction to market devices. *The sociological review*, *55*(2_suppl), 1-12.

Myers, M.D., 2004. Hermeneutics in information systems research. In J. Mingers & L. Willcocks, eds. Social theory and philosophy for information systems. Chichester: Wiley, pp. 103–128

Napoli, P.M., 2011. *Audience evolution: New technologies and the transformation of media audiences*. Columbia University Press.

Napoli, P.M., 2014. The Institutionally Effective Audience in Flux: Social Media and the Reassessment of the Audience Commodity. In: Manzerolle V and McGuigan L (eds) The Audience Commodity in a Digital Age: Revisiting a Critical Theory of Commercial Media. New York: Peter Lang, pp. 115-134.

Negus, K., 2002. The work of cultural intermediaries and the enduring distance between production and consumption. *Cultural studies*, *16*(4), pp.501-515.

Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 1461444818769694.

Papacharissi, Z. (2010). A Networked self: identity, community, and culture on social network sites: Routledge.

Papacharissi, Z. and Easton, E., 2013. In the habitus of the new. *A companion to new media dynamics*, pp.167-184.

Parker, G.G. and Van Alstyne, M.W., 2005. Two-sided network effects: A theory of information product design. *Management science*, *51*(10), pp.1494-1504.

Parker, G.G., Van Alstyne, M.W. and Choudary, S.P., 2016. *Platform revolution: How networked markets are transforming the economy--and how to make them work for you*. WW Norton & Company.

Passoth, J.H., Sutter, T. and Wehner, J., 2014. The Quantified Listener: Reshaping Providers and Audiences with Calculated Measurements.*Mediatized worlds*, pp.271-287.

Patton, M. (1990). Qualitative evaluation and research methods (2nd ed.). Newbury Park, CA: Sage

Patton, M. (1999). Enhancing the quality and credibility of qualitative analysis. *Health services research*, *34*(5 Pt 2), 1189.

Pellikka, H., 2014. 'Gamification in Social Media'. Master's Thesis, University of Oulu Department of Information Processing Science, Finland

Peterson, R.A., 1976. The production of culture: a prolegomenon. *American Behavioral Scientist*, *19*(6), pp.669-684.

Plantin, J.C., Lagoze, C., Edwards, P.N. and Sandvig, C., 2016. Infrastructure studies meet platform studies in the age of Google and Facebook. *new media & society*, p.1461444816661553.

Plantin, J. C. (2018). Review essay: the platform society: public values in a connective world. *Media, Culture & Society*.

Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Potts, Jason, John Hartley, John Banks, Jean Burgess, Rachel Cobcroft, Stuart Cunningham, and Lucy Montgomery (2008), "Consumer Co-creation and Situated Creativity," Industry and Innovation, 15(5), 459-74.

Powers, D., 2015. Intermediaries and intermediation. In: Bennett A and Waksman S (eds) *The Sage Handbook of Popular Music*. Thousand Oaks, California: Sage Publications Ltd, pp.120–134

Puschmann, C., & Burgess, J. (2013). The politics of Twitter data.

Ritzer, G., Dean, P., & Jurgenson, N. (2012). The coming of age of the prosumer. *American behavioral scientist*, *56*(4), 379-398.

Rochet, J.C. and Tirole, J., 2003. Platform competition in two-sided markets. *Journal of the european economic association*, *1*(4), pp.990-1029.

Rosenberg, D. (2013). Data before the fact. In L. Gitelman (Ed.), ''Raw data'' is an oxymoron (pp. 15–40). Cambridge: MIT Press.

Rosenstein, B. (2002). Video use in social science research and program evaluation. *International Journal of Qualitative Methods*, *1*(3), 22-43.

Ruppert, E., Isin, E., & Bigo, D. (2017). Data politics. *Big Data & Society*, 4(2), 1–7.

Santini, R. M. (2018). Recommender systems as "tastemakers": Collaborative filtering as a market strategy for online cultural products. *Observatorio (OBS*)*, *12*(2), 122-135.

Shapiro, C., Carl, S., & Varian, H. R. (1998). *Information rules: a strategic guide to the network economy*. Harvard Business Press.

Sayer, A., 1992. *Method in Social Science. A realist Approach*. London: Routledge.

Sismondo, S., 1993. Some social constructions. *Social studies of Science*,*23*(3), pp.515-553.

Skov, L., 2014. "Cultural Intermediaries and Fashion." In The Cultural Intermediaires Reader, edited by Jennifer Smith Macguire and Toby Miller, 113– 124. London: Sage.

Smith Maguire, J. (2014). Bourdieu on cultural intermediaries. *Smith-Maguire, Jennifer; Mattews, Julian, The cultural intermediaries reader. London: Sage*, 15-24.

Smith-Maguire, J. & Matthews, J. 2012. Are we all cultural intermediaries now? An introduction to cultural intermediaries in context. *European Journal of Cultural Studies* 15: 551–562

Smith-Maguire, J. S., & Matthews, J. (Eds.). (2014). *The cultural intermediaries reader*. Sage.

Smythe, D. W., 1981. On the audience commodity and its work. In: Durham MG and Kellner D (eds) Media and Cultural Studies: Keyworks. Oxford: Blackwell, 2001 ed., pp.253–279.

Stanoevska-Slabeva, K., Sacco, V. and Giardina, M., 2012. Content Curation: a new form of gatewatching for social media. *Documento electrónico. Recuperado el*, *16*.

Staykova, K.S. and Damsgaard, J., 2015, May. A Typology of Multi-sided Platforms: The Core and the Periphery. In *ECIS*.

Tarde, G., 1903. The laws of Imitation, translated by Elsie C. *Parsons, New York: Henry Holt*.

Tilson, D., Lyytinen, K. and Sørensen, C., 2010. Research commentary—Digital infrastructures: The missing IS research agenda. *Information systems research*, *21*(4), pp.748-759.

Tiwana, A. (2013). *Platform ecosystems: aligning architecture, governance, and strategy*. Newnes.

Tiwana A (2014) Platform Ecosystems: Aligning Architecture, Governance, and Strategy. Amsterdam: Morgan Kaufmann Publishers.

Tiwana, A. (2015). Evolutionary competition in platform ecosystems. *Information Systems Research*, *26*(2), 266-281.

Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Research commentary—Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. *Information systems research*, *21*(4), 675-687.

Thorson, K. and Wells, C., 2016. Curated flows: A framework for mapping media exposure in the digital age. *Communication Theory*, *26*(3), pp.309-328.

Vanderbilt, T. (2016). *You May Also Like: Taste in an Age of Endless Choice*. Simon and Schuster.

van Dijck, J., 2013. *The culture of connectivity: A critical history of social media*. Oxford University Press.

van Dijck, J., 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), p.197.

van Dijck, J. and Poell, T., 2013. Understanding social media logic. *Media and Communication*, *1*(1), 2-14.

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: public values in a connective world*. Oxford University Press.

Vargo, S.L. and Lusch, R.F., 2004. Evolving to a new dominant logic for marketing. *Journal of marketing*, *68*(1), pp.1-17.

Veblen, T., 1899. *The theory of the leisure class: An economic study in the evolution of institutions*. Macmillan.

Villi, M., 2012. Social curation in audience communities: UDC (user-distributed content) in the networked media ecosystem. *Participations: The International Journal of Audience and Reception Studies*, *9*(2), pp.614-632.

Webster, J., Gibbins, N., Halford, S., & Hracs, B. J. (2016, May). Towards a theoretical approach for analysing music recommender systems as sociotechnical cultural intermediaries. In *Proceedings of the 8th ACM Conference on Web Science* (pp. 137-145). ACM.

Wegner, P., 1997. Why interaction is more powerful than algorithms. *Communications of the ACM*, *40*(5), pp.80-91.

Winner, L. (1980). Do artifacts have politics?. *Daedalus*, 121-136.

Wright, D., 2015. *Understanding cultural taste: Sensation, skill and sensibility*. Springer.

Yin, R.K., 2003. *Case study research: design and methods.* 3rd ed. Thousand Oaks: Sage Publications.

Yoo, Y., Henfridsson, O. and Lyytinen, K., 2010. Research commentary—the new organizing logic of digital innovation: an agenda for information systems research. *Information systems research*, *21*(4), pp.724-735.

Yoo, Y., Boland Jr, R.J., Lyytinen, K. and Majchrzak, A., 2012. Organizing for innovation in the digitized world. *Organization Science*, *23*(5), pp.1398-1408.

Zajc, M. (2015). Social media, prosumption, and dispositives: New mechanisms of the construction of subjectivity. *Journal of Consumer Culture*, *15*(1), 28-47.

Zhao, X., Salehi, N., Naranjit, S., Alwaalan, S., Voida, S., & Cosley, D. (2013, April). The many faces of Facebook: Experiencing social media as performance, exhibition, and personal archive. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1-10). ACM.