



Follow the algorithm: An exploratory investigation of music on YouTube



Massimo Airoidi^{a,*}, Davide Beraldo^b, Alessandro Gandini^c

^a Department of Social and Political Sciences, University of Milan, via Conservatorio 7, 20122, Milan, Italy

^b Department of Sociology, University of Amsterdam, Nieuwe Achtergrach 166, 1018WV, Amsterdam, Netherlands

^c Media Department, Middlesex University, London, The Burroughs, Hendon, London NW4 4BT, United Kingdom

ARTICLE INFO

Article history:

Received 28 July 2015

Received in revised form 29 April 2016

Accepted 2 May 2016

Available online 13 May 2016

Keywords:

YouTube

Music

Algorithms

Genre

Network analysis

Digital methods

ABSTRACT

This article presents an exploratory study of the network of associations among 22,141 YouTube music videos retrieved by ‘following’ the platform’s recommender algorithm, which automatically suggests a list of ‘related videos’ to the user in response to the video currently being viewed. As YouTube’s recommendations are predominantly based on users’ aggregated practices of sequential viewing, this study aims to inductively reconstruct the resulting associations between the musical content in order to investigate their underlying meanings. Network analysis detects 50 clusters of tightly connected videos characterised by a strong internal homogeneity across different axes of similarity. We discuss these findings with reference to the literature on music genres and classification, arguing that the emerging clusters can be considered as ‘crowd-generated music categories’. That is, sets of musical content that derive from the repeated, crowd-based actions of sequential viewing by users on YouTube in combination with the platform’s algorithm. Interestingly, 7 out of 50 clusters are characterised by what may be seen as a ‘situational’ culture of music reception by digital audiences. Such culture is not so much founded on music genres as traditionally conceived, but rather on the purposes of reception which are rooted in the context where this takes place.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The diffusion of a variety of Internet sources that allow for widespread access to music has had a significant impact on the cultures and practices of music reception as well as on the relationships among individual listeners, musical content and technology. Arguably, a prominent actor at the heart of this process is YouTube, the popular video streaming host owned by Google that is now a global repository for popular music and the entry point for a vast number of listeners-consumers searching for new music (Cayari, 2011; Thelwall, Sud, & Vis, 2012).

This article offers a contribution to the fields of popular music studies, cultural and media sociology by presenting an exploratory study of the network of associations among 22,141 YouTube music videos, as produced by the platform recommender algorithm (Celma, 2010). The aims of this study are to: a) reconstruct how musical content clusters together; and b) understand the meaning of these associations in order to learn about the aggregated practices of sequential viewing by YouTube users and the cultures of reception that surround such practices.

* Corresponding author. Present Address: via Conservatorio 7, 20122, Milan, Italy.

E-mail addresses: massimo.airoidi@unimi.it (M. Airoidi), d.beraldo@uva.nl (D. Beraldo), a.gandini@mdx.ac.uk (A. Gandini).

This work builds on the premise that the network of ‘related videos’ on YouTube is not merely a computational outcome. On the contrary, existing research demonstrates how it originates from the activity of YouTube users who spontaneously upload, tag and consume digital music – a process that is nevertheless technologically-mediated by the platform. YouTube features an automatic system that recommends ‘related’ videos of potential interest, based on site activity (Davidson, Liebold, & Liu, 2010: 293) – that is, in principle, an attempt at guessing users’ musical tastes. Concretely, a list of related videos is normally shown to the side of any visualised content, and this strongly influences users’ consumption pathways on the platform (Zhou et al., 2010). Despite the fact that the current exact formulation of the algorithm is not publicly known, previous contributions by researchers at Google (owner of YouTube) acknowledge how the set of recommended related videos is shaped by the collective behaviour of users, since the YouTube recommendation system relates similar videos mainly on the basis of the most frequent co-views (Bendersky, Garcia-Pueyo, Harmsen, Josifovsk, & Lepikhin, 2014; Davidson et al., 2010). Thus, since the ‘relatedness’ of YouTube music videos is, in principle, the outcome of users’ aggregated viewing patterns, our research aims to ‘follow’ the platform’s recommender algorithm in order to map and interpret resulting associations between different types of music.

From a theoretical perspective, this article aims to expand the sociological understanding of music classifications and genres. This topic has recently gained significant scholarly attention (see Beer, 2013; Lena & Peterson, 2008; Schmutz, 2009; van Venrooij, 2009; Santoro, 2002). However, it still lacks rigorous empirical investigation (Holt 2007: 8). In the sociological literature, music genres are commonly intended as “systems of orientations, expectations, and conventions that bind together an industry, performers, critics, and fans in making what they identify as a distinctive sort of music” (Lena and Peterson, 2008: 698). Far from being rigid and static, cultural classifications of genres are historical, contextual and fluid artefacts (DiMaggio, 1987; Lena & Peterson, 2008; Middleton, 1990); hence, according to DiMaggio, “the challenge for the sociology of art is to understand the processes by which similarities are perceived and genres enacted” (1987:441). This is particularly true now that online platforms have substantially restructured the ways in which music is distributed, discovered and consumed (Prior, 2014; Tepper & Hargittai, 2009). In a recent paper, sociologist David Beer discussed the role of online tagging, searching and recommendations in shaping contemporary music categorisation mechanisms, arguing that “cultural boundaries now appear more open to rapid and unconstrained drawing as a consequence of the media formats through which they are archived, ordered and obtained” (Beer, 2013: 153). Other scholars have similarly stressed the need for new methodological approaches capable of grasping the moving, fine-grained, relational complexity of music classificatory schemes in an inductive way (e.g., Savage & Gayo, 2011; Rimmer, 2012; van Venrooij, 2009). Online research represents an opportunity to step further in this direction. Moving from a critique to the use of standard qualitative and quantitative methods for the study of everyday forms of art categorisation, Beer (2013) suggests to take advantage of the enormous amount of user-generated data publicly available on social media for the study of ‘ground-up’ music reception patterns. Building on this methodological premise, this paper aims to map the associations between related music videos on YouTube, and to inductively explore music reception and classification among digital audiences.

The research methods used in this paper take inspiration from the ‘digital methods’ approach (Rogers, 2013), according to which we should ‘follow the medium’ to study cultural and social phenomena unfolding on the Web. This approach makes it possible to exploit the categories and procedures that digital platforms ‘natively’ adopt to organise information and structure individual behaviour for research purposes. Consequently, we ‘followed’ YouTube’s related videos to derive a network of associations among music videos.

The empirical part relies on two techniques: network analysis (see Wasserman & Faust, 1994) and content analysis (Krippendorff, 2013). The former is used to analyse the above-mentioned network of associations, aggregating highly connected videos into distinct clusters and applying a so-called ‘community detection’ algorithm.¹ This allows us to analyse a ‘second-order’ network composed of clusters, instead of single videos. The second technique we use, content analysis, examines the occurrence of words in the title of a video, which is entered by users themselves. The interpretation of the most recurrent words in each cluster allows us to infer, in cases of a strong coherence between them, the music genres associated with these clusters ‘from the bottom-up’.

Although the effects of the YouTube algorithm on our results remain factually impossible to check (given the undisclosed nature of the algorithm itself), we present what we believe is a unique overview of digital music reception and platform cultures that is, admittedly, exploratory in nature, but has significant potential for replication.

The harvesting of the associations in our relational dataset reveals the existence of 50 clusters of YouTube music videos. The features of this clustering present significant elements of interest for both cultural sociology and the study of music audiences. We argue that these clusters can be considered *crowd-generated music categories* – that is, sets of musical content that are generated by the repeated, crowd-based individual activity of users on the website (see Striphos, 2015). These sets represent aggregated data on the “perception of style and meaning” (Frith, 1996: 94) by digital music audiences on YouTube, which can be seen as products of crowd-generated principles of similarity.

We will discuss our findings and in particular highlight two elements worthy of closer attention. First, assuming that these clusters actually represent crowd-generated music categories, we show how they partially overlap with existing music

¹ Community detection is a network analysis technique that attempts to identify sub-groups of nodes that are internally denser and externally less dense (see Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). In our case, it converts the full network into a small number of clusters, each one composed by videos that are strongly associated.

genres. Second, 7 out of 50 clusters share the reference to what may be seen as a broadly ‘functional’ or ‘situational’ consumption, in which the salience of the context of reception (e.g. an intimate dinner, a party, relaxing or helping babies sleep) becomes the main cultural aspect. This opens up broader reflections on digital music audiences and music reception, since these crowd-generated music categories seemingly indicate a potential departure from traditional notions of music taste and consumption based on genre affiliations (which are often rooted in particular subcultures). Instead, these categories embrace a culture of music reception based on the ‘situational’ purpose that the music video serves in relation to the context and function of reception.

2. Literature review

2.1. Popular music in the digital world

Music today is largely produced, distributed and consumed via digital means (Suhr, 2009; Tepper & Hargittai, 2009). The music industry arguably experienced the breakthrough of digital technologies into the processes of production and, more recently, of distribution and diffusion of music. Since 1999, when Napster first enabled peer-to-peer exchanges of small-size music files free of charge, up to contemporary systems that permit the streaming of large musical databases at a price scale that goes from low to no cost, listening to music has become an intangible, digitally-mediated and increasingly mobile consumption practice (Prior, 2014).

A leading role in this context is played by the video streaming platform YouTube. Founded in 2005 as an independent video-sharing website, it was purchased by Google in 2006 and quickly established itself as one of the giants of the social media industry. According to official sources, YouTube had more than 1 billion monthly users in 2015, and an estimated 300 h of video were uploaded every minute.² Contributions in the literature by Thelwall et al. (2012), Cayari (2011), Green and Burgess (2009) all clearly demonstrate how central music is on YouTube and to the user-created content that is predominant on the platform. In the music industry the traditional broadcast mode of communication in which radio stations and MTV were the main players has now substantially been replaced by an online-based industry. This is shaped by the socio-technical characteristics of digital platforms – potentially giving rise to new forms of reception and cultural understandings of musical content. Music videos remain crucial promotional tools for artists and their branding strategies (Vernallis, 2010), but they are now principally shared on platforms like YouTube and Vevo for purposes of viral diffusion over social media. Arguably, YouTube is the centre of this dynamic as a free-to-use, video-centred ecosystem where user-created content (Suhr, 2009; van Dijk, 2009) blends with official music videos. This unprecedented availability of music content in different formats, potentially impacts not only music production, distribution and reception, but also music discovery, exploration and reception (Baek, 2015; Tepper & Hargittai, 2009). As Rimmer points out, “there now exist a range of interactive resources through which Internet users may become (digitally) converted to new or other musical forms” (2012:303).

Furthermore, this dynamic introduces a crucial interplay with non-human actors and, in particular, recommendation algorithms (see Hallinan & Striphas, 2014; Striphas, 2015). These algorithms feature prominently on platforms such as Amazon, Last.Fm, YouTube and Spotify and play a key role in shaping contemporary music reception and exploration pathways (Celma, 2010), because they provide users with automated suggestions that influence “various decision-making processes, such as what items to buy, what music to listen to, or what online news to read” (Ricci et al., 2011: 1). Flourishing in the e-commerce sector around the early 2000s (Bolton, Katok, & Ockenfels, 2004), recommendation systems work through a variety of methods (see Celma, 2010), normally aggregating similar items and/or users to suggest consumption choices and patterns (Ricci et al., 2011). There is a growing literature in the social sciences discussing the pervasive and invisible power of algorithms over the everyday lives and experiences of Internet users (e.g. Beer, 2009) as well as, more broadly, of citizens (e.g. Cheney-Lippold, 2011). However, so far only few contributions have dealt with the way recommendation systems blend with user interaction (e.g. Striphas, 2015). Despite being a popular topic in computer science, it has to date almost exclusively been addressed from the point of view of performance evaluation (e.g. Celma, 2010; Ricci et al., 2011) and individual user behaviour (e.g. Bolton et al., 2004).

2.2. Music reception and recommendation on YouTube

As mentioned above, YouTube features a recommendation system that automatically produces “a ranked list of related videos shown to the user in response to the video that she is currently viewing” (Bendersky et al., 2014: 1). The YouTube user interface shows by default 25 related videos for any video that is being watched. According to recent publications, the current related video suggestion techniques on YouTube are mainly based on collaborative filtering. The principal data source taken into account by the algorithm are patterns of shared viewership. In other words, if many users watch video A right after video B, these two videos are likely to then be ‘related’ (Bendersky et al., 2014). However, the exact functioning of YouTube’s recommendation system is confidential and – like most proprietary algorithms – it may frequently change (Beer, 2009; Rogers, 2013). Less recent contributions in the literature seem to suggest that YouTube’s recommendation system applies a co-view-based ‘behavioural’ logic (Baluja et al., 2008) or one that is ‘syntactical’, based on matching keywords within the

² Statistics available at <https://www.youtube.com/yt/press/statistics.html> [Accessed: 18 May 2015].

title, description and tags (Cheng, Dale, & Liu, 2008: 236). Either way, as Google researchers Davidson and colleagues write, it seems fair to assume that the YouTube related videos algorithm is principally rooted upon “co-visitation counts” (i.e., co-views); additional data sources such as “time stamps of video watches” and other metadata are believed to be employed only in order to reduce “presentation biases” and “noise” in the resulting list of recommendations (Davidson et al., 2010: 294).

This essentially means two things: a) YouTube videos can be seen as nodes in a network, with related videos inducing a directed graph in which an edge can be established between each pair of videos³ (Davidson et al., 2010: 295); and, b) the weight of the edges is determined by the users' aggregated consumption practices on the platform (Bendersky et al., 2014). The action of co-viewing two media contents is shaped by technical elements, such as the presence of related videos themselves, as well as by the uploaders' activity on the site. This may consist of curating YouTube channels, creating and naming playlists, discursively framing a song as ‘relaxing’ – thus enabling specific keyword-based exploration paths – or managing promotional activities by labels, concert venues, radio shows as well as ‘YouTube stars’ (see van Dijck, 2009).

Up to now, despite the growing importance of Internet research (Coleman, 2010; Marres, 2012; Rogers, 2013), there have been very few studies in the social sciences that directly tackle online streaming services (e.g. Zhang et al., 2013). YouTube has been widely investigated through qualitative and quantitative approaches (see Giglietto, Rossi, & Bennato, 2012), but so far without considering the network of related videos. An interesting study of music categories on recommender websites Audioscrobbler.com and Musicmobs.com was conducted by Lambiotte and Ausloos (2005). The authors examined the large graph linking the users of the websites with online shared music libraries in order to “probe the reality of the usual music divisions, e.g. rock, alternative & punk, classical”, to propose a “quantitative way to define more refined musical subdivisions [. . .] that are not based upon usual standards but rather upon the intrinsic structure of the audience” (Lambiotte & Ausloos, 2005: 2). They detected clusters of artists sharing the same audience, thereby showing that many of these “islands” in the network corresponded to “standard, homogeneous style groupings” (Lambiotte & Ausloos, 2005: 6), such as country, dance, pop, swing, jazz, rock. Other clusters were “geographically localised”, and some revealed “unexpected collective listening habits, thereby uncovering trends in music” (Lambiotte & Ausloos, 2005). However, the authors assumed in principle the existence of a general correspondence between taste clusters and music categories – which has been increasingly questioned by omnivore-univore theorists in the sociology of taste (see Peterson, 2005). Furthermore, they focused on niche music publics – e.g. Radiohead was found to be the group with the largest audience (Lambiotte & Ausloos, 2005) – which makes their results not very generalizable to mainstream audiences.

Our study of YouTube aims to overcome such limitations, since the associations among musical videos on YouTube are not determined by the existence of communities of users owning the same songs in their libraries, but by a multitude of micro-social practices – such as, building playlists and subscribing to channels – that generate aggregated co-viewing patterns. Also, YouTube's worldwide ubiquity allows us to cover both mainstream as well as niche audiences. At the same time, we must consider that music on YouTube is rarely detached from video, and often the visual component is predominant over audio (Holt, 2011).

Existing research has documented how music preferences are influenced by a variety of stimuli that include recommendation algorithms but largely remain a social construction. Tepper and Hargittai (2009) find in their study of music exploration habits by American college students that traditional social circles and mainstream media continue to be important means through which students learn about new music (2009:245). Arguably, it is not just technology that influences music reception, but it could also be the other way around. That is, most recommendation systems relate and distribute content by translating the behaviour of users into automated suggestions (Celma, 2010). Our research aims to expand the current knowledge on this process, and the implications this has for music classification.

3. DATA & METHODS

This study is based on a mixed method approach that combines network analysis with content analysis taken together in a research design inspired by the ‘digital methods’ approach (Rogers, 2013). This latter approach proposes to take advantage of the way digital platforms produce and organise data in order to inform research around the analysis of large scale digital networks. This can be summarized by the motto ‘follow the medium’, which means to “natively” research the digital environment chosen, following the existing logic of action on the platform. In coherence with this methodological framework, we scraped, crawled and analysed a set of YouTube videos in order to reconstruct the network resulting from the aggregated consumption practices of millions of digital music consumers (Davidson et al., 2010).

The data were retrieved using YouTube Data API v. 2.0,⁴ pulled on 26 February 2014. Data collection consisted of two phases. The first phase allowed us to obtain a generic list of videos related to music content, querying the API for the keyword ‘music’ and setting the language parameter to English. The extreme generality of the keyword allows for the inclusion of a

³ A network (or graph) is a collection of nodes (or vertices), which are the items being connected; edges (or links) are the connections between these items.

⁴ An API (Application Programming Interface) is a set of methods used to programmatically access a system. In this case, YouTube Data API allowed us to query YouTube's servers for data related to music videos (a list of videos from a keyword and a list of related videos for each video). YouTube Data API version 2.0 has now been replaced by version 3.0. The documentation is still available here: https://developers.google.com/youtube/2.0/developers_guide_protocol_api_query_parameters [Accessed: 10 January 2016].

wide range of music genres, while the decision to restrict the results to the English language is suggested by the fact that the search term itself is in English. This initial set consists of 500 videos. Note that this is the maximum number of items that the YouTube API v. 2.0 allows one to retrieve when searching videos by keyword.⁵ From this list, we retained only those videos classified as ‘music’ content according to the related parameter included in the meta-data. This way we obtained a ‘seeding sample’ of 333 music videos. In the second phase, we ‘crawled’ the YouTube related videos algorithm in order to build a more consistent sample and to derive a relational dataset for the network analysis. For each video in the seeding sample, we queried the API for the 25 related videos that it allows one to collect, thus expanding our data with new videos and the links among them. After further filtering non-music content, we reiterated the last procedure on the video list we had just obtained. Therefore, we ended up with a network consisting of 22,141 nodes, linked by 70,582 edges.⁶ This procedure is essentially an iterative aggregation of a set of ego-networks, which recalls well-established snowball sampling techniques in classical social network analysis (Browne, 2005). Although the sample does not allow for the generalisation of quantitative findings – such as the distribution of categories – it is nevertheless adequate in providing an exploratory account of how music videos are matched according to the logic of the platform. This is particularly pertinent given the fact that the seeding sample of 333 music videos, which served as a basis to collect the larger dataset, has proven to be heterogeneous in terms of content according to preliminary inspection.

In order to analyse the data, we first applied network analysis techniques with the purpose of exploring the relations that exist among music videos in the sample. The concept of ‘network’ is increasingly adopted in the study of popular music (e.g. Holt, 2007; Webb, 2007), and, consequently, network analysis is also employed more often in this field (e.g. Crossley, 2008). Taking inspiration from existing research that looked at community detection on social media platforms (e.g. Van Meeteren, Poorthuis, & Dugundi, 2010), we concentrated on the so-called ‘community structure’ of the network. More concretely, we applied a community detection algorithm⁷ in order to identify internally connected clusters of videos within the overall network.⁸ This allowed us to disentangle the various bottom-up associations of music content, which are in part related to the spontaneous behaviour of music consumers.

Next, we conducted a content analysis of video titles in order to identify and code the most recurrent words in each cluster (see Krippendorff, 2013) to interpret different cluster associations from a qualitative perspective. In practical terms, we classified the clusters according to the distribution of the observed keywords. This allowed us to explore the interplay between the platform’s affordances and users’ aggregated practices in defining the associational logic underlying the related music videos.

With regards to potential ethical issues concerning user privacy in the collection of digital data (see Boyd & Crawford, 2012), we started from the assumption that the content at our disposal is public, and, moreover, focused on YouTube item-based network and metadata instead of specific users and their activities. This allowed us to avoid potential issues of privacy and ethics, since individual user data are not included in this study.

4. Results

This section is organised as follows. First, we show how the videos cluster in a relatively small number of aggregations. To make sense of these clusters, we inductively describe and name each cluster by analysing the 10 words most frequently used in the titles and the most central content for each cluster. Second, we position the clusters on a bi-dimensional map on the basis of the different categories of words occurring in the video titles. These are coded according to their reference to a plurality of dimensions (all of which are defined below): ‘reception’, ‘milieu’, ‘venue/radio/label’, ‘song/album/tour’, ‘genre’, ‘artist’ and ‘cross-genre’. Finally, we present the networked associations among different clusters.

4.1. Exploring the network of related music videos

The overall network is made of 22,141 videos uploaded by 9994 users. Videos are linked by 70,582 relational ties that are drawn whenever two videos are related by the algorithm (see Fig. 1). The graph is weighted on the basis of connections between pairs of videos occurring more than once. Weights between ties range from 1 to 85, signalling a redundancy of connections between videos, which gives substantial meaning to the analysis of their clustering. Although the relations generated by the algorithm are not necessarily symmetrical, we decided not to treat the graph as a directed network, since we are interested in the overall logic of aggregation (and particularly in cluster detection), rather than the specific structure of the network itself. Therefore, adding directionality to the ties would unnecessarily complicate the analysis.

⁵ For a discussion on the limits of YouTube Data API v. 2.0, see: <https://code.google.com/p/gdata-issues/issues/detail?id=428> [Accessed: 27 December 2015].

⁶ It is important to stress that by using the YouTube APIs, we did not incur in biases due to the machine’s search history on the platform that may impact on the algorithmic selection of recommended videos (Bendersky et al., 2014: 2), since data collection does not require any sort of authentication method.

⁷ More specifically, the Louvain method (Blondel et al., 2008).

⁸ There are different terms to refer to the components resulting from a community detection, like clusters, modules and communities. We decided on ‘cluster’ because the term ‘module’ has quite a different connotation, while the term community would be too ambiguous in a sociological paper. It is important to stress that this analysis has nothing to do with the popular statistical method of cluster analysis, nor with other network analysis measures such as the clustering coefficient.

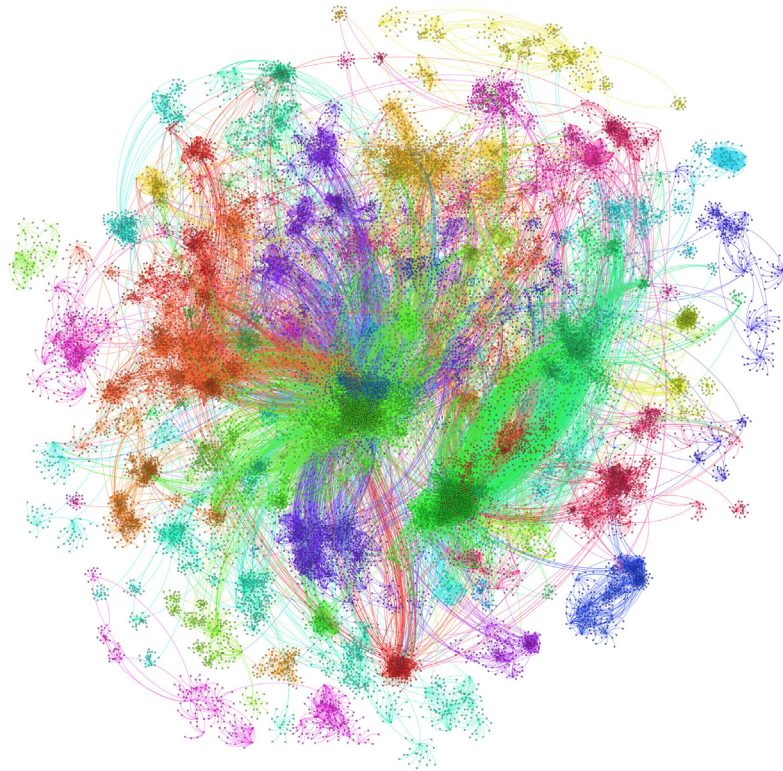


Fig. 1. YouTube related-music videos network. Colours indicate different clusters.

We ran a community detection algorithm (see [Van Meeteren et al., 2010](#)) in order to identify clusters of videos internally connected with each other within the overall network. This procedure results in a ‘modularity’ score, which can be interpreted as the indicator for the ‘goodness of fit’ of this decomposition ([Newman, 2006](#)). We note this is quite high (0.792) and therefore interpret this as a legitimization of the results of the community detection ([Gaul & Klages, 2013](#)). The number of detected clusters is 50, and these clusters are marked using the different colours in [Fig. 1](#). Inspection of this visualisation suggests a fairly high level of aggregation, considering that the dataset is composed of more than 20,000 videos. Furthermore, the size of the majority of clusters is approximately of the same magnitude. This provides additional legitimization to our overall interpretation.

After having detected 50 groups of music videos that have been co-viewed by a vast number of YouTube users, we moved to the second phase of our data analysis in which we investigated the presence of common traits shared by music videos belonging to the same cluster. In order to do so, we first analysed the ‘semantic core’ of each cluster – that is, the ten most frequent words in the video titles, excluding irrelevant ‘stop-words’ (see [Krippendorff, 2013](#)). We did this by labelling each cluster on the basis of the main commonalities shared by its videos, which inductively emerged from the textual content of the titles. Note that this normally summarises the content of the video with a few keywords in order for it to be easily searchable through simple queries. For instance; the ten most recurrent words in Cluster 0 (2049 videos overall) are ‘cover’; ‘One Direction’; ‘live’; ‘lyrics’; ‘Daft Punk’; ‘Ariana Grande’; ‘Katy Perry’; ‘Justin Bieber’; ‘Pharrel Williams’ and ‘Miley Cyrus’. Overall; 48% of the videos in Cluster 0 feature at least one of these words in the title. Among these videos; 170 are cover versions of contemporary international pop songs; while the large majority is constituted by official videos of commonly recognised ‘mainstream’ pop artists. The most viewed content of this cluster is Miley Cyrus’ ‘Wrecking Ball’; while the most connected (that is; the one featuring the highest number of ties; particularly central in the cluster’s graph) is ‘Gorilla’ by Bruno Mars. Thus; we labelled this cluster ‘Teen Pop’ – not because all its music videos can be considered exactly so; but simply because this definition fits most of its content (e.g. One Direction; Ariana Grande; Justin Bieber). The genre labels employed to name the clusters are either derived directly from video titles or; especially in the case of clusters where artist names are more common than genre names; are attached by the authors according to conventional music classifications. Nevertheless; the primary purpose of using such labels is to provide the reader with a rough idea of what these clusters are made up of. They therefore should not be taken as comprehensive categorisations ([Table 1](#); below).

Clusters are internally homogeneous from a semantic point of view. In 27 clusters out of 50, the ‘semantic core’ (i.e. the ten most recurrent relevant words in the titles) is shared by the totality of videos. Also, in the specific case of ‘Sounds of Nature’ the internal semantic relatedness is particularly strong, since 4 out of 10 of the most recurrent words are present in

Table 1

The 50 clusters of YouTube music videos.

Cluster label	N. videos	Average View Count	Video with highest number of ties (title)	% of videos sharing semantic core
Pop Hits	1095	57,881,137	Rihanna – Diamonds	41.5
00s Pop Stars/Latin Pop	117	23,515,812	"Takin' Back My Love" – Enrique Iglesias feat. Ciara	100.0
Teen Pop	2049	15,933,419	Bruno Mars – Gorilla [Official Music Video]	48.1
Rock 90s	1294	11,070,770	Guns N' Roses – November Rain	58.7
80s Pop/Rock	744	7,950,758	Michael Jackson – Thriller	65.6
YouTube Stars	227	6,295,894	Thrift Shop – Macklemore & Ryan Lewis ft Wanz (Alex G Acoustic Cover) Music Video	100.0
Hip Hop	332	5,690,143	Eminem – When I'm Gone	100.0
Piano/Violin Cover	303	5,524,832	Radioactive – Lindsey Stirling and Pentatonix (Imagine Dragons Cover)	98.3
Female Rock Bands/Live	340	5,467,791	Lana Del Rey – Born To Die	91.2
Live Pop	213	4,614,317	Bruno Mars "Locked Out Of Heaven"/SiriusXM/Hits 1	100.0
Gangsta Rap	1023	4,347,026	50Cent – We Up (Explicit) ft. Kendrick Lamar	68.4
Teen Pop Fandom	12	3,822,487	One Direction Funny Moments, Dancing, Lyric Changes & More!	100.0
Arabic Pop	255	3,416,347	Babylone – Zina – Lyrics	72.2
Country	1012	2,798,329	CMA Country Music Festival 2013	59.0
Indie/Alternative Rock	857	2,667,606	The Black Keys – Tighten Up [Official Music Video]	96.7
Best Songs/Top Hits	387	2,284,593	Hit Songs of 2000	100.0
House/Lounge	753	2,011,772	Lounge Beats 2 by Paulo Arruda Deep & Jazz	100.0
Music for Babies	205	1,900,946	🎵🎵 4HOURS of BRAHMS LULLABY 🎵🎵 – BABY SLEEP MUSIC – BABY RELAXING MUSIC – BEDTIME	100.0
Trap Music	628	1,849,700	Trap Music Mix 2013 – October Festival Trap Music Mix (Rave Mix) DJ R3Z	66.7
Soul/Singers/Orchestra	460	1,651,978	James Last & Orchester – Medley 2012	72.6
Dance/Trance	760	1,587,890	Dimitri Vegas & Like Mike – Bringing Home The Madness 21-12-2013 (FULL HD 2HOUR LIVESET)	77.2
Gospel/Worship Music	643	1,409,554	50 Sucessos da Musica Evangelica Atual ***Sucessos da Musica Gospel 2013 e 2014***	82.6
Acoustic Pop Rock/Live	239	1,300,260	Adele: NPR Music Tiny Desk Concert	100.0
Metalcore/Post-Hardcore	578	1,213,398	Hands Like Houses – Introduced Species (Music Video)	86.5
Glee Music	137	1,085,738	GLEE – I Have Nothing (Full Performance) (Official Music Video) HD	100.0
Irish Music	158	1,041,849	Traditional Irish Music – Brogan's Bar – Ennis, Ireland	100.0
Sounds of Nature	11	1,021,172	Thunderstorm and Rain Sounds over the Ocean, 10 h [Sleep Music]	100.0
Latin/World Music/Meditation	486	745,992	Mariachi Los Camperos- Popurri	50.2
Relaxing Background Music	1593	702,072	Background Music Instrumentals – relaxdaily – B-Sides N°1	100.0
Alternative 80s/90s	408	655,269	Depeche mode – Music for the Masses [Album]	100.0
Epic Music/Soundtrack	1269	543,230	The Best of Epic Music 2012–20 tracks – 1-h Full Cinematic EpicMusicVn	79.4
Disco Polo/Disco Music	449	428,059	DISCO POLO ☆ STYCZEŃ – LUTY ☆ 2014 NOWOŚCI!!!	67.3
Hair Dryer Sound	17	409,392	Relaxing Hair Dryer Sound. 2h ASMR	100.0
Female Pop Stars	155	391,519	Lady Gaga – Olympus ft. David Guetta	100.0
Polka	358	376,161	"Two Sisters Polka": Official Music Video	71.8
Flute/Piano Cover/Tutorial	216	309,474	[Flute] Just Give Me A Reason – P!nk ft. Nate Ruess	100.0
New Age/Mantra	45	286,842	I am the light of my soul ~ Sirgun Kaur & Sat Darshan Singh	100.0
Guitar Tutorial/for Musicians	394	277,597	Mozart Documentary – The Genius of Mozart 1/3 "Miracle of Nature"	85.3
Jazz/Classical	18	275,343	Jacques Loussier Play Bach Trio – Jazzwoche Burghausen 2007	100.0
Soundtrack/Classical	345	258,339	Xenoblade Chronicles Music – Colony 9	82.6
Metal/Rock Live/Documentary	616	243,534	Megadeth-The Story Behind The Music (FULL Documentary)	72.1
Ambient/Chillout	91	234,571	Pleiadian Friend (ambient/psychill/psybient mix)	100.0
Jazz/Bebop	21	185,948	Dizzy Gillespie – Jivin' In Be Bop	100.0
Ethiopia/South Sudan Music	301	118,036	Best New Ethiopian Music 2014 Emebet Negasi Min Yishalal	81.4
Ugandan Music	358	116,705	Dance Non stop music Video Mix 2013 Afroberliner HD	100.0
Pop Stars Interviews	13	101,171	Lady Gaga: Interview (Alan Carr Chatty Man 2013) Pt. 1	100.0
K-Pop	68	81,244	EXO & Apink	100.0
Afro Celt	22	38,389	Afro Celt Sound System – Magic & Mayhem	100.0
Grime/Uk Hip Hop	26	22,482	The Grime Show: Big H	100.0
Celtic Music	40	19,524	Eleanor McEvoy Celtic Connections 2011	100.0
Tot or Average	22 141	3,723,528		87.5

the title of each video. Most of the tightly connected groups of videos detected by our network analysis refer to conventional music genres (e.g. 'Polka'), local music scenes (e.g. 'Grime/Uk Hip Hop') and generational music preferences (e.g. 'Rock 90s'). Yet, some of these clusters also seem to point at different semantic dimensions that depart from conventional definitions of music genre (e.g. 'Music for Babies').

4.2. Music clusters as crowd-generated music categories

In order to analyse this discrepancy between cluster types in more detail, we compared the musical content in each cluster by coding the ten most frequent relevant words in the title of the videos. This way, we scrutinize the underlying patterns of semantic association between different musical contents that are implicitly emerging from the network structure and stem from users' aggregated practices. Building on Webb's idea of *milieu* (2007), that encapsulates geographical, cultural and social aspects of music scenes, we coded with the tag 'milieu' all words explicitly referring to a local context of music production (e.g. 'South Sudan music', 'Irish music'). With the tag 'reception' we coded all words that refer to the context of music reception (e.g. 'sleep', 'relax'). The tag 'genre' was used for words pointing to conventional music genres, while 'cross-genre' classified common words generically referring to the features of a music video (e.g. 'live', 'remix', 'instrumental'). Furthermore, the tag 'venue/radio/label' indicates words that relate to the world of music making (music venues, radio shows, music labels), and the tag 'Youtuber/other media' refers to the so-called 'YouTube stars' (cf. musicians and video-makers who are popular because of their videos on YouTube) and media platforms such as iTunes or Spotify. The remaining categories (artist, song/album/tour) are self-evident.

There is evidence that clusters differ considerably in relation to the relative frequency of the categories of words described above. We see that several clusters are characterised by an extremely high rate of videos containing artists' names and conventional genre names, and relatively few common names that are related to the musical content, while many other clusters display an opposite trend. For instance, in 'Ambient/Chillout', 'Jazz/Classical', 'House/Lounge' and 'Jazz/Bebop' more than 90% of the video titles share the same references to genres and artists, whereas less than 50% include generic words referring to the content of the videos. At the same time, more than 90% of the videos belonging to clusters such as 'Sounds of Nature', 'Hair Dryer Sound', 'Glee Music' and 'Best Songs/Top Hits' are described by 'cross-genre' words in the titles, and fewer than 25% by shared references to specific artists and music genres. We therefore observe an analytical polarity in the axis between those clusters mainly connoted for 'genre' and those mainly connoted by 'cross-genre'. Another similar polarity emerges if we pay attention to how more than 60% of the videos included in 'Ugandan Music', 'Pop Stars Interviews', 'Grime/UK Hip Hop' and 'Irish Music' are discursively framed via a reference to a common cultural milieu (Webb, 2007) and/or to specific music venues, radios or music labels (which are key elements in any field of music production, see Bourdieu, 1993). These very same clusters stand out by the absence of references to the contexts of music reception. On the contrary, for clusters in which the category 'reception' is significantly important – such as in 'Guitar Tutorial/For Musicians' (21.8%), 'Relaxing Background Music' (60.6%), 'Hair Dryer Sound' (100%), 'Sounds of Nature' (100%), 'Music for Babies' (100%) – the categories 'milieu' and 'venue/radio/label' never appear. A second semantic axis is thus established based on the prominence of production or reception-related coding. See Table 1 for examples of video titles per cluster.

In order to better interpret and compare the positioning of each cluster according to these two dimensions – and following Middleton's suggestion to "locate musical categories topographically" (1990:7) – we constructed a Cartesian graph and assigned x and y coordinates to each cluster. The x-axis corresponds to the 'genre'/'cross-genre' continuum. For each cluster we computed the offset between the relative frequency of cross-genre labels on the one hand, and artist and genre words on the other. We aggregated 'artist' and 'genre' labels because both semantically refer to the actual musical content of the video, implicitly or explicitly qualified in terms of genre. The y-axis analogously refers to the offset between the percentage of words in the title related to the context of reception and those related to the context of production. Here, we aggregated the categories 'milieu' and 'venue/radio/label' for purposes similar to those above (Fig. 2).⁹

On the left side of the map, among those clusters sharing genre-identity of sorts (negative x-values), one can see the opposition between the negative y-values of clusters close to classical definitions of music genre (e.g. 'Ugandan Music', 'Grime/UK Hip Hop', 'Irish Music', 'Ethiopia/South Sudan Music', 'K-Pop', 'Celtic Music', 'Jazz/Bebop') and the positive y-values of clusters closer to newer genres (such as 'House/Lounge', 'Ambient/Chillout', 'Pop Hits', 'Rock 90s', '80s Pop/Rock'). Whereas the former can be described as a "constellation of styles connected by a sense of tradition", the second show resemblances to "marketing categories" instead (see Holt, 2007: 18).

On the right side, ($x > 0$; 'cross-genre' polarity), conventional genres are replaced by less common music categories, such as 'Sound of Nature', 'Hair Dryer Sound', 'Best Songs/Top Hits', 'Glee Music', 'YouTube Stars' and 'Flute/Piano Cover/Tutorial'. All these clusters refer to videos explicitly sharing cross-genre commonalities, such as the 'relaxing' tones of a thunderstorm, the sound of a fan or a hairdryer, merely being a part of a 'top 10 songs' ranking or of the soundtrack of a television series such as 'Glee', the strong presence of online user-generated contents (van Dijck, 2009) and/or of specific musical instruments.

The highest y-values – which denote the discursive prevalence of the contexts of music reception over those of music production – are associated with 'Music for Babies' ($y = 3$), 'Sounds of Nature' ($y = 2.7$), 'Hair Dryer Sound' ($y = 2.7$), 'Relaxing

⁹ These are the adopted formulas: $x = f_{\text{cross-genre}} - (f_{\text{artist}} + f_{\text{genre}})$; $y = f_{\text{reception}} - (f_{\text{milieu}} + f_{\text{venue/radio/label}})$. Values on the map are normalised according to axes' means and standard deviations.

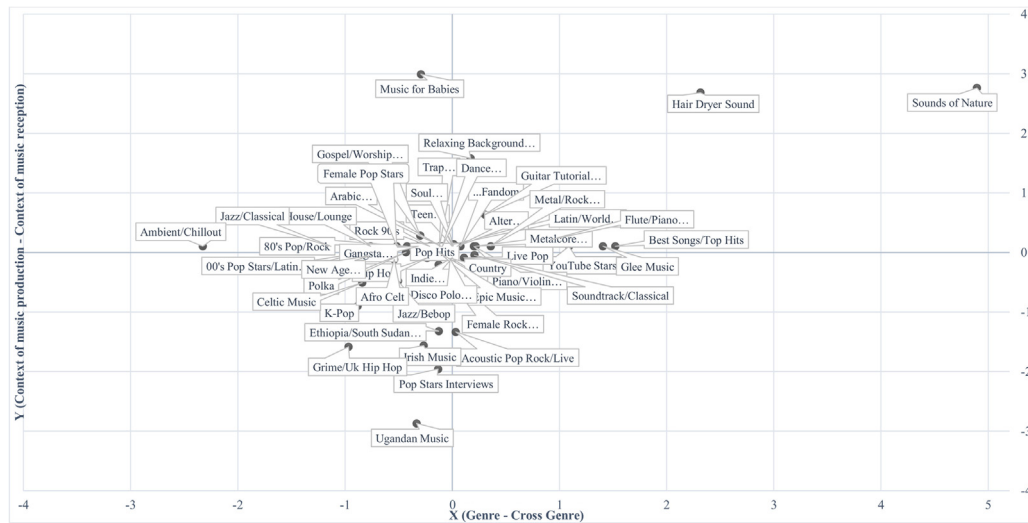


Fig. 2. Music clusters on the semantic map, positioned according to the types of words occurring in video titles.

Background Music' ($y = 1.6$), 'Guitar Tutorial/For Musicians' ($y = 0.6$), 'Gospel/Worship Music' ($y = 0.3$), 'Flute/Piano Cover/Tutorial' ($y = 0.2$). In all these clusters, what is particularly evident is the salience of the 'situational purpose' of the music (or sounds). 'Sounds of Nature' and 'Hair Dryer Sound' music videos are always portrayed 'for relaxing' or 'for sleeping'; the same (plus 'meditation') applies to 'Relaxing Background Music'. 'Music for Babies' and 'Gospel/Worship Music', respectively, include music for keeping babies calm and for religious worship, whilst the remaining two clusters are collections of tutorials and videos 'for musicians'. The higher the y -value, the more references to the context of music reception in the titles of the videos, and the more a specific word becomes frequent: 'hour(s)'. Therefore, the value of this 'situational music' seems to depend more on its duration – which of course has to be long enough to musically support what listeners are doing at the time – than on its perceived quality or on the reputation of the performers. Finally, 8 out of 50 clusters are particularly close to the origin of the axes: 'Teen Pop', 'Pop Hits', 'Alternative 80s/90s', 'Dance/Trance', 'Soul/Singers/Orchestra', 'Trap Music', 'Teen Pop Fandom' and 'Arabic Pop'. Clusters based on pop Charts – that are normally neither stylistically homogeneous (Lena & Peterson, 2008: 700) nor rooted in a specific 'milieu' (Webb, 2007) of production – all reside in this interval.

Thanks to the semantic map shown above, we have been able to devise an analytic scheme to interpret and compare the different logics of similarity emerging from the videos' discursive frames (Goffman, 1974). Although the binary oppositions 'genre/cross-genre' and 'context of production/context of reception' are not the only suitable heuristic tools for this purpose, they proved very useful to show the semantic discrepancy that we outlined.

Our analysis shows that the retrieved music clusters are not meaningless agglomerations of videos. On the contrary, there seems to exist an underlying cultural logic of similarity within each cluster that is produced by the technologically-mediated and aggregated practices of usages by listeners and uploaders. The social logic of aggregation pertaining to these clusters is particularly evident in the case of 'Ethiopia/South Sudan Music', which includes 301 music videos. The majority of these videos (137) explicitly refer to the Ethiopian musical scene, whereas 26 mention South Sudan and 10 are spontaneously defined as 'Nuer Music'. Nuer are a tribal ethnic group on the border of Ethiopia and South Sudan, also studied by the anthropologist Evans-Pritchard (1940). The strong connections between these three geographically and culturally close traditions are evident in the structure of the related-videos network, which in this case clearly materialises a local musical milieu. The complex interplay between the culturally-informed practices of music makers and listeners acting on the platform (e.g. building a genre-based playlist, promoting an artist's official YouTube channel, uploading one's favourite songs or, simply, listening to music) is translated and simplified by the recommender system's technical logic. The result is that a small cluster in a network of 22,141 nodes reproduces a coherent miniature of a very specific musical culture. There are no Ethiopian songs in more 'western' music clusters or even among Ugandan music videos, and vice versa. In particular, the degree centrality measure of this latter cluster in the network analysis is actually 0. This means that there are no connections with any other cluster. This last example – that is, a culturally and stylistically-coherent cluster composed by music videos defined by various textual frames (Ethiopian music, Nuer music, South Sudan music) – supports our assumption of the predominance of a co-view-based 'behavioural' logic over a 'syntactical' one in the current functioning of the YouTube algorithm (see Section 2.2). Otherwise, the prevalence of a 'natural language processing' approach based on video metadata would have undermined the results of our text analysis.

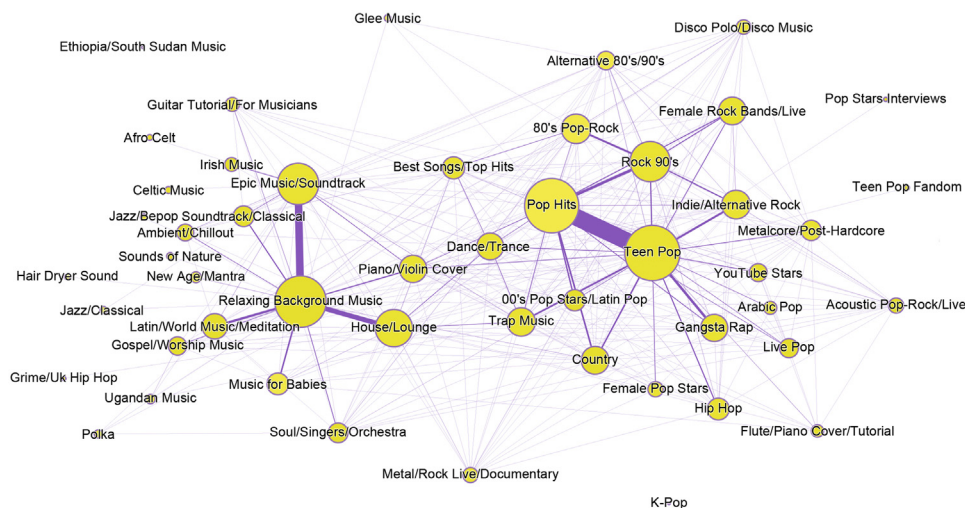


Fig. 3. The network of associations among clusters.

4.3. Inter-cluster associations

To complete our exploration of music on YouTube, we examined the clustered network. This network is derived from the community structure of our original related videos network: nodes are now clusters of strongly related videos, while links among clusters correspond to the sum of the links among videos belonging to different clusters. As a result of our interpretation of the clusters as 'crowd-generated music categories', the weight of an edge linking two clusters can be seen as an indicator of the 'relatedness' between two categories. The clustered network is composed of 50 nodes (the clusters) and 463 edges, as shown in Fig. 3.¹⁰ The shape of the visualised network is computed with Gephi¹¹ using the OpenOrd algorithm, designed to highlight clusters.

The strongest connections in this network tend to link clusters characterised by similar musical content (e.g. 'Pop Hits' with 'Teen Pop'; 'Relaxing Background Music' with 'Epic Music/Soundtrack'). Furthermore, the network is roughly divided into two main 'meta-clusters', one of which mostly contains clusters largely referring to conventional music categories (e.g. pop, rock, country), while the other is characterised by what may be seen as a prevalently 'situational' logic (e.g. relaxing, background, meditation, music for babies).

Since the probability of clusters being connected is affected by the size of the same clusters, Table 2 lists the ten strongest associations between clusters divided by the size of the source cluster. Each row thus provides the strongest relative associations in network. Again, we can conclude that connections between clusters show a substantially high rate of stylistic contiguity.

5. Discussion, digital music between 'genre' and 'situation'

The analysis of the network associations within a large sample of YouTube music videos has shown how the users' aggregated reception practices – which we assumed as determinant in the relatedness between two music videos in combination with the recommender algorithm – produce groupings that can be interpreted as *crowd-generated music categories* (see Striphas, 2015). These groupings are categories that emerge from the interplay of platform "affordances" (Baym & Boyd, 2012) and the reception patterns of music communities (see Fabbri, 1982). Through a content analysis of videos' titles and the interpretation of the resulting 'semantic map', we also provide evidentiary support on how a peculiar 'logic of similarity' seems to characterise these crowd-generated music categories and their content in an ideal continuum that goes – on the one hand – from conventional genre affiliations to the presence of specific cross-genre stylistic traits and – on the other hand – from the semantic prevalence of the context of music production to that of music reception.

The interpretation of our clusters as crowd-generated music categories seems legitimate considering that Franco Fabbri defines music genres as sets "of musical events (real or possible) whose course is governed by a definite set of socially accepted rules", which are shared by a given "community" of listeners (1982:52–53). Similarly, Simon Frith stated that genre "is not determined by the form or style of a text itself but by the audience's perception of its style and meaning" (1996:94). Put differently, in coherence with these authors that support the idea of the eminently social character of music classification

¹⁰ The size of the bubbles is proportional to the number of videos that pertain to the cluster. The thicker the link, the higher the edge weight.

¹¹ See <https://gephi.org/> [Accessed: 12 June 2015].

Table 2
Top 10 inter-cluster associations.

Source	Target	Edge weight/N. videos source
Pop Hits	Teen Pop	94.2
Indie/Alternative Rock	Teen Pop	49.3
House/Lounge	Relaxing Background Music	45.0
Soundtrack/Classical	Relaxing Background Music	24.7
Rock 90s	Pop Hits	20.4
00s Pop Stars/Latin Pop	Pop Hits	17.1
Country	Pop Hits	9.5
Female Rock Bands/Live	Rock 90s	9.1
Live Pop	Pop Hits	7.4
Relaxing Background Music	Epic Music/Soundtrack	6.9

(see also [Lena & Peterson, 2008](#)), we argue that the tightly connected groups of related music videos coming out of the computational analysis of the community structure of the network should be considered as rough miniatures of music categories as they emerge “from the ground-up” ([Beer, 2013:153](#)). Importantly, we note that conventional music genres still appear to be one of the major structuring forces in guiding listeners’ shared reception patterns on the platform.

Interestingly, however, a kind of ‘situational’ or ‘functional’ consumption also emerges, whereby the listener sets apart purely stylistic or aesthetic conceptions of music genres to choose their soundtrack based on the effect it has on daily activities. Existing research regards this as typical of a ‘popular’ approach to music that is distant from the formal intellectual appreciation characterising ‘high art’ (see [Frith, 1996](#); [van Venrooij & Schmutz, 2010](#)). “Functional” (or “*umgangsmäßig*”) genres such as dance, entertainment and liturgical music have been traditionally considered by musicologists as “trivial” or spurious, being just “one partial aspect of an event that is determined by extra-musical factors” (see [Dahlhaus, 2004: 263](#)). Nevertheless, we show how the reference to a ‘situational purpose of music reception’ characterises 7 out of 50 clusters, which feature music pieces tailored for relaxation, meditation, religious worship, and so on. Although this is not exactly a new trend in contemporary music markets – see, for instance, electronic music producer Brian Eno’s “Music for Airports”, DeNora’s considerations about music and social situations [DeNora \(2000:11–14\)](#) and Fabbri’s reflections on background listening [Fabbri \(2003\)](#) – we expected it to be relatively marginal. Instead, it seems to be a quite relevant presence if we consider that more than 10% of YouTube music in our sample follows this logic. It is also worth mentioning that a similar ‘situational’ form of categorisation is observed on other online sources of digital music – such as the streaming service Spotify, where ready-made playlists are often characterised by comparable ‘situational’ frames. The acknowledgment of such an emerging trend in current digital music reception seems to be particularly important if we think about the subcultural meaning of music that historically plays a key role in the formation of youth identity and peer recognition in adolescence ([Hall & Jefferson, 1976](#)).

Both the analyses of the clusters’ content and of the resulting network of clustered associations confirm that ‘relatedness’ on YouTube is generally synonymous of ‘stylistic contiguity’. Our results may sound obvious at a superficial glance: being a commercial service provider, YouTube is naturally interested in giving meaningful suggestions to its users ([Celma, 2010](#)). Still, if it is true that the musical field “together with its internal relationships, is never still – it is always in movement” ([Middleton, 1990: 7](#)), then it could be argued that studies like this offer room for tracking its transformations in an inductive way (see also [Beer, 2013](#); [Savage & Gayo, 2011](#); [van Venrooij, 2009](#)).

We acknowledge that our analysis has some limitations, particularly in relation to the fact that our 22,141 videos are not a statistically representative sample of the music content uploaded on YouTube. Yet, it must also be recognised that this would have been a major concern if the goal had been to provide inferences on the composition of the overall repository of YouTube videos related to music, or to perform quantitative comparisons. Instead, for the purposes of the present work, the snowball sampling is coherent with the main aim of the analysis; that is, doing an exploratory study into the logic of association among videos. We hope that future studies may build on our work, and pursue this more ambitious line of inquiry. With regards to the nature of our findings, it must also be acknowledged that some of the clusters are not entirely composed of musical contents (‘Pop Stars Interview’, ‘Metal/Rock Live/Documentary’, ‘Guitar Tutorial/For Musicians’, ‘Flute/Piano Cover/Tutorial’), while others are aesthetically heterogeneous (‘Soul/Singers/Orchestra’, ‘Latin/World Music/Meditation’). More generally speaking, the logic of the recommendation algorithm, primarily based on mass listening behaviour, implies the risk of observing “different cultures of categorization [. . .] removed from their social basis” ([Holt, 2007: 29](#)).

Nevertheless, this article provides further empirical support to the argument that digital platforms offer new epistemological and methodological tools to cultural sociology to investigate social patterns by inductively exploring users’ practices and imaginaries (see [Beer, 2013](#)). More particularly with regards to the sociological study of music, we believe that a network-based approach that is capable of grasping the relational nature of music’s semantic space is probably what is now most useful for the study of genres (as other empirical research has also shown, e.g. [Crossley, 2008](#); [van Venrooij, 2009](#)). Following those contributions that approach music classification in an eminently sociological way, our methodological perspective, if replicated on a larger scale, can arguably shed light on the ‘folk categories’ used by contemporary digital music

listeners to classify musical contents (see Fabbri, 2003), while avoiding the risk of naturalising conventional genre categories (see DiMaggio, 1987: 441).

Acknowledgments

In addition to the Editor and the anonymous referees, we would like to thank Alex van Venrooij and Rens Wilderom for their helpful support with this article. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Baek, Y. M. (2015). Relationship between cultural distance and cross-cultural music video consumption on YouTube. *Social Science Computer Review*, 33(6), 730–748.
- Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., et al. (2008). Video suggestion and discovery for YouTube: taking random walks through the view graph. In *proceedings of the 17th international conference on world wide web* 895–904.
- Baym, N., & Boyd, D. (2012). Socially mediated publicness: an introduction. *Journal of Broadcasting & Electronic Media*, 56(3), 320–329.
- Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, 11(6), 985–1002. <http://dx.doi.org/10.1177/1461444809336551>.
- Beer, D. (2013). Genre, boundary drawing and the classificatory imagination. *Cultural Sociology*, 7(2), 145–160. <http://dx.doi.org/10.1177/1749975512473461>.
- Bendersky, M., Garcia-Pueyo, L., Harmsen, J., Josifovski, V., & Lepikhin, D. (2014). Up Next: Retrieval Methods for Large Scale Related Video Suggestion Categories and Subject Descriptors. *KDD'14*.
- Blondel, V. D., Guillaume, J., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. In *J. Stat. Mech.* 1–12.
- Bolton, G. E., Katok, E., & Ockenfels, A. (2004). How effective are electronic reputation mechanisms? An experimental investigation. *Management Science*, 50(11), 1587–1602.
- Bourdieu, P. (1993). In R. Johnson (Ed.), *The field of cultural production: essays on art and literature* New York: Columbia University Press.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data information. *Communication & Society*, 15(5), 662–679. <http://dx.doi.org/10.1080/1369118X.2012.678878>.
- Browne, K. (2005). Snowball sampling: using social networks to research non-heterosexual women. *International Journal of Social Research Methodology*, 8, 47–70.
- Cayari, C. (2011). The YouTube effect: how YouTube has provided new ways to consume, create, and share music. *International Journal of Education & the Arts*, 12(6), 1–28.
- Celma, O. (2010). *Music recommendation and discovery: the long tail, long fail, and long play in the digital music space*. Heidelberg, Dordrecht, London, New York: Springer.
- Cheney-Lippold, J. (2011). A new algorithmic identity: soft biopolitics and the modulation of control theory. *Culture & Society*, 28(6), 164–181. <http://dx.doi.org/10.1177/0263276411424420>.
- Cheng, X., Dale, C., & Liu, J. (2008). Statistics and social network of youtube videos. *IWQoS*, 2008, 229–238.
- Coleman, E. G. (2010). Ethnographic approaches to digital media. *Annual Review of Anthropology*, 39(1), 487–505. <http://dx.doi.org/10.1146/annurev.anthro.012809.104945>.
- Crossley, N. (2008). Pretty connected: the social network of the early UK punk movement theory. *Culture & Society*, 25(6), 89–116. <http://dx.doi.org/10.1177/0263276408095546>.
- Dahlhaus, C. (2004). Trivial music and aesthetic judgement. In C. J. Washburne, & M. Derno (Eds.), *Bad music: the music we love to hate* (pp. 262–277). New York & London: Routledge.
- Davidson, J., Liebold, B., & Liu, J. (2010). The YouTube video recommendation system. *Proc. ACM RecSys* 293–296.
- DeNora, T. (2000). *Music in everyday life*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511489433>.
- DiMaggio, P. (1987). Classification in art. *American Sociological Review*, 52(4), 440–455.
- Evans-Pritchard, E. E. (1940). *The Nuer*. Oxford: Clarendon Press.
- Fabbri, F. (1982). A theory of musical genres: two applications. In P. Horn, & D. Tagg (Eds.), *Popular music perspectives* (pp. 52–81). Exeter: International Association for the Study of Popular Music.
- Fabbri, F. (2003). Taboo listening. *Background listening and music composition conference*.
- Frith, S. (1996). *Performing rites: on the value of popular music*. Cambridge, MA: Harvard: University Press.
- Gaul, W., & Klages, R. (2013). A hierarchical clustering approach to modularity maximization. In A. Lausen, Berthold, Van den Poel, Dirk, & Ultsch (Eds.), *Algorithms from and for nature and life. studies in classification, data analysis, and knowledge organization* (pp. 79–86). Springer.
- Giglietto, F., Rossi, L., & Bennato, D. (2012). The open laboratory: limits and possibilities of using facebook, twitter, and YouTube as a research data source. *Journal of Technology in Human Services*, 30(3–4), 145–159.
- Goffman, E. (1974). *Frame analysis: an essay of the organization of experience*. New York: Harper and Row.
- Green, J., & Burgess, J. (2009). *YouTube: online video and participatory culture*. Cambridge: Polity Press.
- Hall, S., & Jefferson, T. (1976). *Resistance through rituals: youth subcultures in post-war Britain*. London: Hutchinson.
- Hallinan, B., & Strihas, T. (2014). Recommended for you: the Netflix Prize and the production of algorithmic culture. *New Media & Society*. <http://dx.doi.org/10.1177/1461444814538646>.
- Holt, F. (2007). *Genre in popular music*. Chicago: University of Chicago Press.
- Holt, F. (2011). Is music becoming more visual? Online video content in the music industry. *Visual Studies*, 26(1), 50–61. <http://dx.doi.org/10.1080/1472586X.2011.548489>.
- Krippendorff, K. (2013). *Content analysis: an introduction to its methodology*, 3rd ed. London: Sage.
- Lambiotte, R., & Ausloos, M. (2005). Uncovering collective listening habits and music genres in bipartite networks. *Physical Review E* 72, 72(6), 066107.
- Lena, J., & Peterson, R. (2008). Classification as culture: types and trajectories of music genres. *American Sociological Review*, 73(5), 697–718.
- Marrs, N. (2012). The redistribution of methods: on intervention in digital social research, broadly conceived. *The Sociological Review*, 60, 139–165. <http://dx.doi.org/10.1111/j.1467-954x.2012.02121.x>.
- Middleton, R. (1990). *Studying popular music*. Buckingham and Philadelphia: Open University Press.
- Newman, M. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Science*, 103(23), 8577–8582.
- Peterson, R. A. (2005). Problems in comparative research: the example of omnivorousness. *Poetics*, 33(5–6), 257–282. <http://dx.doi.org/10.1016/j.poetic.2005.10.002>.
- Prior, N. (2014). The plural iPod: a study of technology in action. *Poetics*, 42, 22–39. <http://dx.doi.org/10.1016/j.poetic.2013.11.001>.
- Ricci, F., Rokach, L., & Shapira, B. (2011). In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender systems handbook* Boston, MA: Springer US. http://dx.doi.org/10.1007/978-0-387-85820-3_1.
- Rimmer, M. (2012). Beyond omnivores and univores: the promise of a concept of musical habitus. *Cultural Sociology*, 6(3), 299–318. <http://dx.doi.org/10.1177/1749975511401278>.

- Rogers, R. (2013). *Digital methods*. Boston, MA: MIT Press.
- Santoro, M. (2002). What is a cantautore?: Distinction and authorship in Italian (popular) music. *Poetics*, 30, 111–132.
- Savage, M., & Gayo, M. (2011). Unravelling the omnivore: a field analysis of contemporary musical taste in the United Kingdom. *Poetics*, 39(5), 337–357. <http://dx.doi.org/10.1016/j.poetic.2011.07.001>.
- Schmutz, V. (2009). Social and symbolic boundaries in newspaper coverage of music, 1955–2005: Gender and genre in the US, France, Germany, and the Netherlands. *Poetics*, 37(4), 298–314. <http://dx.doi.org/10.1016/j.poetic.2009.06.006>.
- Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, 18(4–5), 395–412. <http://dx.doi.org/10.1177/1367549415577392>.
- Suhr, H. C. (2009). Underpinning the paradoxes in the artistic fields of MySpace: the problematization of values and popularity in convergence culture. *New Media & Society*, 11(1–2), 179–198. <http://dx.doi.org/10.1177/1461444808100158>.
- Tepper, S. J., & Hargittai, E. (2009). Pathways to music exploration in a digital age. *Poetics*, 37(3), 227–249. <http://dx.doi.org/10.1016/j.poetic.2009.03.003>.
- Thelwall, M., Sud, P., & Vis, F. (2012). Commenting on YouTube Videos: from Guatemalan Rock to El Big Bang. *Journal of the American Society for Information Science and Technology*, 63(3), 616–629. <http://dx.doi.org/10.1002/asi>.
- van Dijck, J. (2009). Users like you? Theorizing agency in user-generated content Media. *Culture & Society*, 31(1), 41–58. <http://dx.doi.org/10.1177/0163443708098245>.
- van Venrooij, A., & Schmutz, V. (2010). The evaluation of popular music in the United States, Germany and the Netherlands: a comparison of the use of high art and popular aesthetic criteria. *Cultural Sociology*, 4(3), 395–421.
- van Venrooij, A. (2009). The aesthetic discourse space of popular music: 1985–86 and 2004–05. *Poetics*, 37(4), 315–332. <http://dx.doi.org/10.1016/j.poetic.2009.06.005>.
- Van Meeteren, M., Poorthuis, A., & Dugundi, E. (2010). Mapping communities in large virtual social networks. *Proceedings of the IEEE Workshop on Business Applications of Social Network* (pp. 1–8).
- Vernallis, C. (2010). Music video and YouTube: new aesthetics and generic transformations case study: beyoncé's and lady gaga's «Telephone». In H. Keazor (Ed.), *Rewind, play, fast forward: the past, present and future of the music video* Bielefeld: Transcript Verlag.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. Cambridge: Cambridge University Press.
- Webb, P. (2007). *Exploring the networked worlds of popular music*. London & New York: Routledge.
- Zhang, B., Kreitz, G., Isaksson, M., Ubillos, J., Urdaneta, G., Pouwelse, J. A., et al. (2013). Understanding user behavior in spotify. *INFOCOM, 2013*, .
- Zhou, R., Khemmarat, S., & Gao, L. (2010). The impact of YouTube recommendation system on video. *IMC'10* (pp. 404–410).

Massimo Airoidi is a PhD candidate in Sociology at the Department of Social and Political Sciences, University of Milan. He collaborates with the Centre for Digital Ethnography and with the Public Opinion & Media Lab, University of Milan. His main research interests are: cultural taste, digital media, online algorithms, text analysis, digital methods, communication research.

Davide Beraldo is completing his PhD as part of a joint degree program between the University of Amsterdam (Dep. of Sociology) and the University of Milan (Dep. of Social and Political Sciences). He collaborates with the DATACTIVE research group (University of Amsterdam) and with the Centre for Digital Ethnography (University of Milan). His main research interests are social movements, digital methods, online network analysis and epistemology of complexity.

Alessandro Gandini is a Lecturer in the Media Department, Middlesex University, London, and is research active at the Centre for Digital Ethnography, University of Milan. His research interests include the transformations of work in digital society, the intermediation of digital technologies within the processes of value production and distribution, and the research approaches to study digital environments and interaction.