

*Culture Unbound:*  
*Journal of Current Cultural Research*

Thematic Section:  
Discovering Spotify

Edited by  
Rasmus Fleicher and Pelle Snickars

Extraction from Volume 9, Issue 2 2017

Linköping University Electronic Press  
Culture Unbound: ISSN 2000-1525 (online)  
URL: <http://www.cultureunbound.ep.liu.se/>

© 2017 The Authors.

***Culture Unbound, Extraction from Volume 9, 2017***

***Thematic Section: Discovering Spotify***

**Rasmus Fleischer and Pelle Snickars**

*Discovering Spotify – A Thematic Introduction* ..... 130

**Rasmus Fleischer**

*If the Song has No Price, is it Still a Commodity? Rethinking the  
Commodification of Digital Music* ..... 146

**Maria Eriksson and Anna Johansson**

*Tracking Gendered Streams* ..... 163

**Pelle Snickars**

*More of the Same – On Spotify Radio*..... 184

**Roger Mähler and Patrick Vonderau**

*Studying Ad Targeting with Digital Methods: The Case of Spotify*..... 212



## Discovering Spotify – A Thematic Introduction<sup>i</sup>

By Rasmus Fleischer & Pelle Snickars

With a user base now officially reaching more than 100 million, which includes 60 million paying subscribers, the music streaming platform Spotify is today widely recognized as the solution to problems caused by recent decades of digital disruption within the music and media industries. Spotify resembles Netflix, YouTube, and Apple Music as an epitome of streaming's digital Zeitgeist that is shaping our future. Industry interviews, trade papers, academic books, and the daily press reiterate numerous versions of this “technological solutionism” (Morozov 2013) in almost as many variations.

This thematic section of *Culture Unbound* is broadly concerned with the music service Spotify, and novel ways to situate and do academic research around streaming media. Approached through various forms of digital methods, Spotify serves as the object of study. The four articles presented here—three full length research articles and a shorter reflection—emanates from the cross-disciplinary research project “Streaming Heritage: Following Files in Digital Music Distribution”. It was initially conceived at the National Library of Sweden (hence the heritage connection), but the project has predominantly been located at the Umeå University's digital humanities hub, Humlab, where the research group has continuously worked with the lab's programmers. The project involves four researchers and one PhD student and is funded by the Swedish Research Council between 2014 and 2018.<sup>ii</sup>

While most previous scholarship on Spotify has primarily focused on its service role within the music industry, its alterations to the digital music economy, or its influence on ending music piracy (Wikström 2013, Wikström & DeFilippi 2016, Allen Anderson 2015, Galuszka 2015, Andersson Schwarz 2013), our

Fleischer, Rasmus & Pelle Snickars: “Discovering Spotify—A Thematic Introduction”, *Culture Unbound*, Volume 9, issue 2, 2017: 130–145. Published by Linköping University Electronic Press: <http://www.cultureunbound.ep.liu.se>

project mainly takes a software studies and digital humanities approach towards streaming media. The project “Streaming Heritage” broadly engages in reverse engineering Spotify’s algorithms, aggregation procedures, metadata, and valuation strategies to study platform logics, including underlying norms and structures. Reverse engineering starts with the final product (in our case the music service Spotify) and tries to take it apart, backwards, step-by-step. Basically, we draw a more holistic picture by using Spotify as a lens to explore social, technical, and economic processes associated with digital media distribution. The key research idea within our project is to follow files (rather than the people making, using, or collecting them) on their distributive journey through the streaming ecosystem, taking empirical advantage of inherent data flows at media platforms (such as Spotify).

Over the last ten years, the extensive field of media and Internet studies have used several digital methods to develop pioneering ways to analyse and understand the digital, the Internet, as well as digital media production, distribution, and consumption. Following the catchphrase “the system is the method” (Bruhn Jensen 2011), digital methodologies are increasingly deployed to perform social science or humanistic inquiries on, for example, big data and black-boxed media platforms (such as Spotify) (Ruppert, Law & Savage 2013). As a research practice, digital methods “strive to follow the evolving methods of the medium” (Rogers 2013:1). The issue of data of, about, and around the Internet, as Klaus Bruhn Jensen has eloquently stated, “highlights the common distinction between research evidence that is either ‘found’ or ‘made’”. If one disregards various complexities, basically all evidence needed for Internet or digital studies is already at hand. When interacting, searching, and listening to music at Spotify, for example, user data are constantly being produced. Such data are “documented in and of the system” and “with a little help from network administrators and service providers” it can be used as the empirical base for research (Bruhn Jensen 2011:52).

For researchers seeking to take empirical advantage of data flows at contemporary media platforms, it quickly becomes apparent “that such platforms do not present us with raw data, but rather with specially formatted information” (Marres & Gerlitz 2015). Data, in short, are often biased. Twitter, for example, determines what data are available and how the data can be accessed, and researchers often have a hard time knowing what relevant data might be missing. Hence, the major academic problem confronting media scholars working with digital methods is the lack of access to data. In our project, the main difficulty in doing research on and around Spotify is the reluctance of the company to share data.

Consequently, user data must be acquired and compiled through other means such as by deploying bots as research informants or by recording and aggregating self-produced music and sounds. Building on the tradition of breaching ex-

periments in ethnomethodology (Garfinkel 1967), where reactions are caused by disturbing or even violating commonly accepted rules or norms, our project has tried via repeated and modified so-called “interventions” to break into the hidden infrastructures of digital music distribution. On the one hand, we have been interested in broadly studying different data patterns and media processes at Spotify. On the other hand, we have also been keen on producing and obtaining research data, for example, by using bots as virtual listeners, by documenting (and tracing) Spotify’s history through constantly changing interfaces, or by tracking and archiving advertisement flows. Using debugging software such as Fiddler or Ghostery, we have also tracked traffic between a computer and the Internet.

Although this thematic section of *Culture Unbound* is concerned with Spotify, basically any other streaming media services could be studied in similar ways. The various digital methods we present, use, and critically discuss can be used to analyse a range of different online services or platforms that today serve as key delivery mechanisms for works of culture, including YouTube, Netflix as well as various platforms for e-books or academic articles. Although our analysis is specific, the methods we propose are of more general relevance and concern. For example, using bots as research informants can be deployed for many different types of digital scholarship. Due to the transformation of media into data, digital methods can easily be used in research (albeit with some coding skills). When media at online services (such as Spotify) are coded and redefined as a purely data-driven communication form—with, on the one hand, content (e.g., media files and metadata) being aggregated through external intermediaries, and, on the other hand, user-generated data being extracted from listening habits—the singularity of the media experience is transformed and blended into what Jeremy Wade Morris has termed “a multimediated computing experience” (Wade Morris 2015: 191).

For a regular user, today’s multimediated and exceedingly computational experience of online media takes on different and sometimes personalised forms. To understand the logic and rationale of contemporary media services and platforms, one should not shy away from but rather ask what exactly happens when data are turned into media and vice versa. What occurs and takes place beneath the black shiny surface of, say, the Spotify desktop client, with its green and greyish interface details and whited fonts and textures? It goes without saying, that research on the cultural implications of software—whether in the form of software studies, digital humanities, platform studies, or media archaeology—has repeatedly stressed the need for in-depth investigations on how computing technologies work combined with (more or less) meticulous descriptions of technical specificities (Kirschenbaum 2008, Chun 2011, Sterne, 2012, Ernst 2013).

## Localising Spotify

Departing from the interventionist and experimental approaches we have used in our research project, which both metaphorically and practically try to track and follow the transformation of audio files into streamed experiences in the simple way a postman would follow the route of a parcel from packaging to delivery, the notion of localisation has become salient. Following files is a technical impossibility in a streaming media context, yet approaching, encircling, and circumscribing Spotify, both as a company and a service, has also proven to be hard. In our research project, we have repeatedly asked insidiously simple questions: Where is Spotify? When is Spotify used? What is Spotify? It might seem naive, but during the research process it has become increasingly difficult for us to understand and grasp our object of study.

Asking Google the search question “What is a Spotify?” returns a snippet from Wikipedia: “Spotify is a music, podcast, and video streaming service, officially launched on 7 October 2008. It is developed by start-up Spotify AB in Stockholm, Sweden” (Wikipedia 2017). But such an answer hides more than it shows and can easily be problematized. Is Spotify, for example, a content platform, a distribution service, or a media company? Furthermore, music naturally lies at the heart of Spotify (even if podcasts and videos seem increasingly important), but what kind of content is accepted—i.e., how is music defined? And what about the Swedishness of Spotify? Where is the company located? Headquarters are still to be found in central Stockholm on Birger Jarlsgatan 61, but the service is now available in some 60 countries, not to mention the digital variety of desktop and mobile versions (which all differ slightly). In addition, how does one situate Spotify commercially and financially (i.e., how much money is Spotify making (or losing) and how can one measure its economic impact?

As is apparent from the four issues above—and one could easily have included yet another—localising Spotify is easier said than done. Starting, however, by determining whether Spotify is a tech or a media company, it was obvious that Spotify for several years foremost offered a technological solution for record companies struggling with piracy. In a private conversation in 2012, one of the authors of this introduction (Snickars) asked Sophia Bendz (at the time Head of Marketing at Spotify) what kind of company Spotify actually was. Without hesitating, Bendz stated that Spotify was a tech company, only distributing content produced by others. The tech identity, however, was somewhat dubious even in 2012 and has become increasingly harder to sustain. Advertisement serves an illustrative case in point. In endless discussions with record labels (around rights management), Spotify took the stance that the continuous offering of a zero-price version with recurrent advertisement (Spotify Free) would in the long run be the best solution, as this strategy would serve as an incentive to scale businesses and attract glo-

bal listeners. Spotify's classification as strictly a tech company misses the fact that a core part of its business has been to provide content to audiences and selling those audiences to advertisers. Other music streaming services used a different strategy and Spotify has consequently struggled, and increasingly become more of a media company, all in order to keep to its business model with two versions of the product: the Free version (with embedded advertising) and the Premium version (without advertising).

Arguably, the music industry still sees Spotify as the top streaming service around, yet Spotify "has done little to address the lack of new music from a large collection of major artists when their albums are released" (Singleton 2016). That is, in a digital environment where streaming music becomes default, a focus on tech and distribution will only result in missed business opportunities. Indeed, Spotify has not really entered into content production (e.g., like Netflix), although some self-made videos are provided such as interviews with artists as well as other content (e.g., pop-ups that explain lyrics). Hence, stating that Spotify is only a tech company (in the form of a streaming service) fails to see other defining characteristics of the enterprise.

Secondly, "Music for everyone" is the company catch phrase, displayed, for example, when entering [spotify.com](https://spotify.com). To localise Spotify, one might ask what kind of music does the service offer? In fact, one fundamental question we have struggled with in our research project is determining what sounds are perceived as music according to Spotify. It should be stressed that uploading music onto the service is outsourced to several so-called aggregation services. In short, these (and not Spotify) regulate content appearing on different music streaming platforms. In one of our interventions, we experimented with uploading self-produced music via different aggregators. These explorations with artificial sounds and music resulted in different responses. The same music (or sounds) passed some aggregators, but others did not define these "sounds" as music content at all. In short, rejection criteria of music aggregators turned out to be arbitrary. Hence, when principles as to what is considered music vary at the aggregation level, and consequently on streaming platforms such as Spotify, usually depending on whether users pay an aggregation fee or not, the line between music and non-music, artist and machine, becomes increasingly blurred.

A third way to use the notion of localisation to pinpoint Spotify is to look closer at geography and the hype around the "Swedishness" of Spotify. On the one hand, the company is still often associated with Sweden: "Swedish music-streaming service provider Spotify is in advanced talks to acquire German rival SoundCloud" (*The Guardian*, 2016). Yet, on the other hand, geographical localisation strategies also make it apparent that Spotify tries hard to transform itself into a global media company: "Spotify is tailoring its service for local tastes, from topical playlists to

tiered pricing, as it prepares to expand its music streaming in Asia” (*Bloomberg* 2016). Spotify, in fact, increasingly acts as a global media company, and as a result, Patrick Vonderau (one of the researchers in our project) has recently claimed that “Spotify is neither particularly Swedish nor about music”. While invocations of the company’s Swedishness have been needed to sustain venture capital, and a “vision of ‘European unicorns’ . . . to position Spotify at the sexy, cool end of digital innovation”, Vonderau argues that in financial terms Spotify now acts more “as a digital broker whose history of equity rounds, market and debt capitalization, and board of directors firmly ties brokerage strategies to U.S.-based financial interests” (Vonderau 2017). Spotify, in short, operates increasingly like a traditional American media company.

A fourth way to try to frame and localise Spotify is to follow the money and look at the company’s evasive finances. Some figures estimate that the company makes more than two billion dollars a year from subscription fees and advertising, yet approximately 80 percent of that income is (all likely) paid to record labels and artists. In general, the financial situation and status of Spotify remains concealed, yet the same basically goes for the commodity that is being sold. As Rasmus Fleischer argues in his article in this thematic section, a crucial issue when dealing with the political economy of digital media is understanding what kind of commodity is being sold and to whom.

Lately, it has even been claimed that Spotify is “causing a major problem for economists” (Edwards 2016). Within mainstream economics, it is now commonly acknowledged that GDP is just an empirical construct that is becoming ever more misleading (Coyle 2014, Economist 2016). One main problem is how to measure inflation: to establish a price index, it is necessary to quantify differences in quality between last year’s products and this year’s products. It is difficult to compare the price of music sold as discrete units and music bundled as a monthly subscription (Spotify Premium) or offered with advertisements (Spotify Free). Is it meaningful to calculate a hypothetical “price per track listened to” in any of these cases? And how should we measure, in monetary terms, the value of music recommendations? Because of such quandaries, economists like Erik Brynjolfsson and Andrew McAfee have pointed to Spotify as an example of how national accounts fail to capture the “consumer surplus” resulting from rapid technological progress (Brynjolfsson & McAfee 2014: 174–189). Even a more traditional calculation of national accounts, which only includes those transactions where money is changing hands, poses delicate problems when locating Spotify. Thus, recent government inquiries from Sweden and the U.K. have singled out Spotify as the epitome of problems with measuring an economy increasingly built on digital services (Felländer 2015; Bean 2016). It seems that Spotify has not only disrupted the music and media industries but also has disrupted the ways in which the economic sta-



tistics surrounding user data need to be measured and interpreted.

## Historicising Spotify

The story of Spotify is commonly told as an extraordinary success story: over 100 million users and over \$8 billion valuation and growing. However, Spotify has yet to show a profit. So far, its losses have tended to grow faster than its turnover, so the survival of the service depends on ever larger injections of venture capital. This situation, typical for today's technology start-ups, tends to limit the opportunities for independent research. To attract investment and to secure deals with partner companies, it is necessary for Spotify to maintain a certain level of buzz in the news media, confirming the image of a company always expanding, always innovating, and always headed on a straight path towards a future monopoly position. No information will be let out if it does not play a predefined role in this public relations strategy.

One might argue that the buzz and hype, including problems in localising the company, makes it difficult for researchers to approach Spotify, at least compared to more established companies that have already gone public. Throughout most of Spotify's lifetime, there have been speculations about an imminent stock market launch, an IPO (Initial Public Offering), or a possible acquisition in which Spotify would be bought up by Google, Apple, or Facebook. Certain commentators have also questioned whether Spotify's business model is sustainable. These discussions and speculations have not lead anywhere and often remain obscure as vital details are kept secret via nondisclosure agreements between Spotify and the music industry. Another impossible (but lively) discussion has been concerned with whether Spotify is good for artists, as if artists exist as a homogenous group to which Spotify can be either good or bad.

From our research perspective, it is more relevant to ask how Spotify takes part in a redefinition of what it means to be a successful artist or a record company by changing the ways in which music is presented, commodified, and valued. In other words, the producer of musical recordings cannot be thought of as existing independently of the distributor. As researchers, we must simply acknowledge that Spotify is a moving object and that the results from our digital experiments and interventions must be situated within a historical context (even though the company is not much older than ten years). One important source material for the historiography of Spotify, which has been essential for our research, is a major archive of news reports, including trade journals focusing on tech (e.g., *Wired* and *Techcrunch*), music (e.g., *Billboard* and *Music Week*) and advertising (e.g., *Advertising Age* and *Marketing Week*), all sources we have constantly been collecting.

Going through this archive, one is confronted by an immense level of buzz,

speculations, rumours, and empty promises. Localising and historicising Spotify is in many ways a task of how one approaches this constant murmur. One possibility is to regard this buzz simply as a kind of noise that ought to be filtered out, leaving a smaller selection of verified stories, useful for producing a historiography over what Spotify has really done. We propose the opposite approach, however: Just as we follow the files using digital methods, we follow the buzz using archives (i.e., our historiography). This means working through a tremendous source material looking not only for what happened, but also after what Richard Barbrook has described as “the beta version of a science fiction dream: the imaginary future” (2007). The history of Spotify is, in fact, full of false predictions and visions. Taking these shortcomings into account provides an important corrective to the conventional narrative about the gradual realisation of a grandiose entrepreneurial vision.

It may surely be true that Spotify CEO, Daniel Ek, has a deep passion for music and that he enjoys playing the guitar, but when he and Martin Lorentzon founded Spotify in 2006, it was certainly not an attempt to disrupt the music industry to save it from piracy, as the official story now goes (Bertoni 2012). The original idea behind Spotify was purely technological: to create a platform for media distribution based on a peer-to-peer network. The first news reports in Sweden, in fact, presented Spotify as a company building a new infrastructure for film distribution. However, because video demanded too much bandwidth, Spotify’s first set up and trials used music files as distribution content (Åkesson 2007, Johansson 2015). To be more precise, the beta version of Spotify was loaded with pirated music files, downloaded by its employees through file-sharing services like The Pirate Bay (Andersson Schwarz 2013: 149). Music streaming proved attractive, and soon enough Ek and Lorentzon had conceived a business model for music, clearly inspired by the popularity of illicit file-sharing in Sweden. Spotify was to make music free but legal, available to consumers at no cost, while advertising provided all revenues.

Spotify’s launch, thus, coincided perfectly with the broader hype around the idea that “\$0.00 Is the Future of Business” (Anderson 2008, Fleischer 2017), but also with the onset of a global financial crisis, which was soon to decimate the advertising market, making it hard to sustain ad-funded “free” services. The business of selling subscriptions for media services, however, tended to do remarkably well in the recession (Economist 2009). Spotify hence gradually changed its mind, now declaring that both advertising and subscriptions were to be equally important sides of their business model, while also dabbling with ideas of making money on sales of merchandise and concert tickets. In retrospect, it is striking how long the founders of Spotify resisted the idea of building a business fully dependent on subscription revenues.

Historiography cannot do without an element of periodisation. With respect to Spotify's financial uncertainty and its dependence on venture capital, the company history can thus be understood over a timeline of investments. These have come in a series of funding rounds, from the first round (Series A) of about \$20 million to the most recent round of \$1 billion in convertible debt. Each time, the value of existing stocks has been diluted, the balance of ownership displaced in a new direction. The identity of the investors is usually public information, aggregated on websites like Crunchbase (2016), but the conditions detailed in each deal is always a secret. However, if one follows the buzz and maps it over the investment timeline, some of it becomes evident. Investments have, for example, been used mostly for international expansion (Series D, Series F) or for developing the streaming service in a specific direction (Series E).

Daniel Ek has been dubbed "the most important man in music" by *Forbes* (Bertoni 2012) and one of the ten most powerful people in the music industry (*Billboard* 2016), yet he is not in control of Spotify. The company's founders most certainly lost their majority share by 2009. In addition, Spotify's existence remains dependent on the willingness of the Big Three record labels (Universal Music Group, Sony Music Entertainment, and Warner Music Group) to renew their licensing deals. Hence there are several reasons why Spotify is not like Facebook: it is not profitable, it is not publicly traded, and it cannot dictate the terms in dealing with content providers. It would be silly to deny that Spotify is not dominant and mighty, but the power of Spotify is not easily located. Rather than being a single forceful actor trying to shape the future of music, Spotify indeed exists at the intersection of competing industries (tech, content, advertising, and finance).

One way to historicise Spotify in a more concrete manner is to look at altered strategies for music discovery. In the earlier period before its U.S. launch, Spotify's interface was centred around the search box (Fleischer 2015). Not much effort was put into assisting users who did not immediately know what music they wanted to hear. In other words, Spotify's ideal user was an individual with strong musical preferences (as part of his or her identity). When asked about the lack of social features in 2009, a Spotify director simply answered: "We're coming at it from the on-demand side" (*Music Week* 2009). This was also Spotify's real strength, according to influential magazines like *Billboard* and *Wired*; the service was considered fast, clean, and easy to use, and importantly so because it did not push music recommendations to its users (Bruno 2009, Peoples 2010, Pollack 2011).

This partly began to change in 2010–11, when Spotify established a strategic partnership with Facebook, following a Series C investment by Sean Parker (co-founder of Facebook and, before that, of Napster) who also joined Spotify's board of directors. The interface was gradually redesigned, moving away from the individualism of the search box and towards more social approaches of friction-

less sharing: all music listening would be automatically shared with friends. This was met, however, by an outcry from many users, forcing Spotify to introduce new options for protecting the privacy of musical preferences (Spotify 2011a, Spotify 2011b, *Financial Times* 2011). In short, the social turn provided a new direction for Spotify's developers, moving away from the poverty of the empty search box and towards a third way, different from both algorithmic and expert-curated music recommendations (Fleischer 2017). By integrating with Facebook, Spotify hoped to create the ultimate discovery engine. Spotify's approach was to recommend music based on what the user's friends had put in their playlists. Friends, however, can have bad taste. Ultimately, social discovery turned out to be a failure in the light of Spotify's experience on the U.S. market. Spotify had emphasised the freedom to choose, but many Americans seemed to prefer the freedom from choice. By the end of 2012, Daniel Ek admitted that "Spotify is great when you know what music you want to listen to, but not so great when you don't" (Bercovici 2012).

Spotify's social turn was followed, just a couple of years later, by a curatorial turn. The development of this type of new music discovery approach (throughout 2013) was financed by a \$100 million investment round (series E) led by Goldman Sachs. Spotify was indeed not a vanguard in this movement. During 2012, industry observers began establishing as a fact that people love to simply lean back and listen. The future of streaming music was now more commonly sought in radio-like lean-back services such as Pandora, while the lean-forward approach of Spotify was seen as its Achilles' heel (Peoples 2012, Warren 2012). Trying to remedy this, Spotify first acquired Tunigo, a company specialised in building expert-curated music playlists. At the same time, Spotify discarded its old, individualist slogan: "Whatever you want, whenever you want it" (Spotify 2011c). New slogans were put in use: "Music for every moment" (Spotify 2013a) and "Soundtrack your life" (Spotify 2013b). In every country where Spotify was active, the local office began to recruit playlist curators with knowledge of local culture, but not specialised in any specific genres. The standard job description used was typical of Spotify's new approach: playlist curators should identify "songs to fit different situations" and create "playlist listening experiences for a multitude of moods, moments, and genres" (Spotify 2014). Here, it seems that Spotify had opted for a more human approach of expert curation, but Spotify was simultaneously working on algorithmic recommendation systems in close cooperation with the music intelligence company The Echo Nest, which it acquired in 2014. Neither a purely human nor a purely algorithmic curation system would be conceivable, but a combination of the two could work. In any case, it is finally interesting to note how this dichotomy was reinforced in 2015 by Apple when it presented its new streaming music service. Apple Music was then framed as the more warm and human alternative to the allegedly cold and all-too-algorithmic Spotify (Apple 2015, Dredge 2015).

## About the Articles

As is evident from the discussions above, analysing Spotify is not an easy task. If localising Spotify is hard, historicizing the company's whereabouts doesn't result in a particularly straight trajectory either. On the contrary, users, competitors, and investors have all influenced the different directions that Spotify has taken and will all likely continue to do so. Hence, if music discovery today is important for Spotify to both satisfy and create a desire to consume and listen to more music, discovering Spotify is another matter. This thematic section of *Culture Unbound*, however, tries to locate the streaming service from several different perspectives. It brings together ongoing and differentiated research within the project "Streaming Heritage: Following Files in Digital Music Distribution". The four articles presented are, in short, all concerned with uncovering and finding out more about Spotify via different research strategies and methods. Three of the articles use digital methods in their approach, trying to get closer to Spotify through inventive experiments. Two of the longer articles (Eriksson & Johansson and Snickars) also explicitly use bots as research informants. A bot is a small software application that runs automated tasks (or scripts), and within interventions at Humlab we have repeatedly used massive set-ups of bots, sometimes working with up to 500 virtual listeners.

In the first article in the thematic section, "If the song has no price, is it still a commodity?", Rasmus Fleischer reviews some of the recent literature on how music is marketed. Over the last century, music has been subject to different regimes of commodification, sold as a published score, as a live performance, or as recorded sound. Streaming services like Spotify, however, represent a different commodification regime, Fleischer argues. Therefore, it is necessary to identify and define the commodity Spotify sells. Fleischer criticises prevalent conceptions of the digital music commodity that often assume that each song (whether downloaded or streamed) is a commodity, which is indeed correct in the case of downloading services like the iTunes Store. But the user of Spotify will (currently) never see a price tag on a song. In fact, Spotify is not selling discrete pieces of recorded sound and is not offering consumers millions of commodities; Spotify offer only one commodity: the subscription. This product is a bundle that includes not only access to all songs in the catalogue, but also the maintenance of a personalised profile connected to a variety of playlists tailored for pre-defined activities. Music is still commodified by Spotify, Fleischer argues, but as a commodity, music can mean different things. Spotify is, for example, buying music through various aggregation services in the form of copyright licenses, bundling it, adding new features, and then selling music as a personalised experience. When analysing commodification, it is always necessary to ask what kind of object is the commodity.

In their article, “Tracking Gendered Streams”, Maria Eriksson and Anna Johansson investigate whether music recommendations at Spotify are gendered. As is well known, one of the most prominent features on contemporary music services is the provision of personalised music recommendations that come about through the profiling of users and audiences. Based on a range of bot experiments, their article explores patterns in music recommendations provided by Spotify in its Discover feature. The article specifically focuses on issues around gender and explores whether the Spotify client and its music recommendation algorithms are performative of gendered user identities and taste constellations. Exploring the tension between gendered publics and Spotify’s promise to deliver personalised music recommendations to everyone, Eriksson and Johansson’s research ties into broader questions about the workings and effects of algorithmic knowledge production. They argue that issues around gender are important in this context, since Spotify’s music recommendations can be considered as one of the venues where gendered norms and ideals are reproduced and manifested. Eriksson and Johansson’s results for example reveal that male artists were highly overrepresented in Spotify’s music recommendations; an issue which they argue prompts users to reproduce hegemonic masculine norms within the music industries. Although the results should be approached as highly historically and contextually contingent, Eriksson and Johansson argue that they do give some evidence of the ways in which gender becomes tied to issues of taste and identity formation in algorithmic knowledge-making processes.

In his article, “More of the Same – On Spotify Radio”, Pelle Snickars takes a similar approach as Eriksson and Johansson, working extensively with bots as research informants. Snickars main interest is the so-called radio function at streaming services, and Spotify Radio in particular. It is a service that “lets you sit back and listen to music you love. The more you personalise the stations to match your tastes the better they get”, at least according to the company slogan. Basically, the radio functionality allows users (via various unknown algorithms) to find new music within Spotify’s vast back-catalogue, offering a potential infinite avenue of discovery. Nevertheless, the radio service has also been disliked and blamed for playing the same artists over and over. Together with the Humlab programmers, Snickars set up an experiment to explore the possible limitations and restraints found within “infinite archives” of music streaming services. The hypothesis was that the radio function of Spotify does not consist of an infinite series of songs although it may appear so to the listener; it is actually a finite loop. Spotify Radio claims to be personalised and never-ending, yet music seems to be delivered in limited loop patterns. What would such loop patterns look like? The intervention used 160 bot listeners programmed to listen to different Swedish music from the 1970s. Snickars is not primarily interested in personalised recommendations,

but rather how Spotify Radio functions generically. The first (and major) round of bots started Spotify Radio based on the highly popular Abba song “Dancing Queen” (with some 65 million streams). The second (and minor) round of bots used the less well-known Swedish progressive rock band Råg i Ryggen’s “Queen of Darkness” (with some 10,000 streams). Snickars article describes different research strategies when dealing with proprietary data as well as the background and the establishment of the radio functionality at streaming services like Spotify. Essentially, his article empirically recounts, discusses, and analyses the radio looping interventions set up at Humlab.

Finally, in their co-written article, “Studying Ad Targeting with Digital Methods: The Case of Spotify”, Patrick Vonderau and Roger Mähler provide a brief description of digital methods used in studying digital advertising technologies. To study ad targeting, researchers have an inventory of tested methods at their disposal but a problem of access to verifiable data persists. In order to understand which types of key stakeholders are involved in ad targeting processes, the authors experimented with digital tools to complement data collection. In doing so, they followed the well-established idea of taking up methods that are already embedded in digital infrastructures and practices.

This thematic section of *Culture Unbound* goes under the hood of Spotify and looks critically at its tech stack. It is important to remember that Spotify’s data infrastructure resembles other services. The analyses put forth in the different articles (sometimes) approximates media specific readings of the computational base; that is, the mathematical structures underlying various interfaces and surfaces resonate with media scholarly interests in technically rigorous ways of understanding the operations of material technologies. Then again, it is also important to stress that the Spotify infrastructure is hardly a uniform platform. Rather it is downright traversed by unseen data flows, file transfers, and information retrieval in all kinds of directions, be they metadata traffic identifying music, aggregation of audio content, playout of streaming audio formats (in different quality ratings), programmatic advertising (modelled on finance’s stock exchanges), or interactions with other services (notably social media platforms). This thematic section tries to uncover and make visible some of these streams.

**Rasmus Fleischer** is a postdoctoral researcher based at the Department of Economic History, Stockholm University. His research interests are located in the intersections between culture and economy, as well as technology and politics. Most of all, he has explored 20th century media history, transformations of copyright and the commodification of music. E-mail: [rasmus.fleischer@ekohist.su.se](mailto:rasmus.fleischer@ekohist.su.se)

**Pelle Snickars** is Professor of Media and Communication Studies, specialising in digital humanities at Umeå university and is affiliated with the Humlab research centre. His research focuses on the relationship between old and new media, media economy, digitisation of cultural heritage, media history as well as the importance of new technical infrastructures for the humanities. E-mail: pelle.snickars@umu.se

## Notes

<sup>i</sup> Some of the digital methods used in this thematic section are non-compliant with Spotify's Terms of Service (ToS). The data collection has ended and did not involve any user data. With the public and academic interest in mind, we appreciate Spotify's forbearance with any trespassings of ToS that our data collection involved.

<sup>ii</sup> The research project, "Streaming Heritage. Following Files in Digital Music Distribution" involves system developers Roger Mähler and Johan von Boer (at Humlab, Umeå University), as well as researchers Pelle Snickars, Maria Eriksson, Anna Johansson (at Umeå University), Rasmus Fleischer (at Stockholm and Umeå University), and Patrick Vonderau (at Stockholm University). For more information: <http://streaming-heritage.se/>.

## References

- Åkesson, Nils (2007): "De kan tjäna 900 miljoner", *Dagens Industri*, January 16.
- Allen Anderson, Paul (2015): "Neo-Muzak and the Business of Mood" *Critical Inquiry* 41(4), 811-840.
- Andersson Schwarz, Jonas (2013): *Online File Sharing: Innovations in Media Consumption*, London: Routledge.
- Apple (2015): Video recording from June event 2015, June 8. <http://www.apple.com/live/2015-june-event/>
- Barbrook, Richard (2007): *Imaginary futures: from thinking machines to the global village*, London: Pluto.
- Bean, Charles (2016): "Independent review of UK economic statistics", *HM Treasury Report* <https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-final-report> (Accessed 15/03/17)
- Bercovici, Jeff (2012): "Spotify Adds Features, Taking a Page From, Yes, Myspace", *Forbes*, December 6.
- Bertoni, Steven (2012): "Spotify's Daniel Ek: The Most Important Man In Music", *Forbes* January 4. <http://www.forbes.com/sites/stevenbertoni/2012/01/04/spotify-daniel-ek-the-most-important-man-in-music/#5d77b2520263> (Accessed 15/03/17)
- Billboard* (2016): "Billboard's 2016 Power 100 List Revealed", *Billboard*, February 12. <http://www.billboard.com/articles/business/6874816/billboards-2016-power-100-list-reveal> (Accessed 15/03/17)
- Bloomberg* (2016): "Spotify focuses on localisation to smooth next expansion in Asia", July 20. <http://www.businesstimes.com.sg/consumer/spotify-focuses-on-localisa->



- [tion-to-smooth-next-expansion-in-asia](#) (Accessed 15/03/17)
- Bruno, Antony (2009): "See Spotify run", *Billboard*, 121(25), p. 11
- Bruhn Jensen, Klaus (2011): "New Media, Old Methods—Internet Methodologies and the Online/Offline Divide", Mia Consalvo & Charles Ess (eds): *The Handbook of Internet Studies*, London: Blackwell.
- Brynjolfsson, Erik & Andrew McAfee (2014): *The second machine age: work, progress, and prosperity in a time of brilliant technologies*, New York: W. W. Norton & Company.
- Chun, Wendy (2011): *Programmed Visions: Software and Memory*, Cambridge, Mass: MIT Press.
- Coyle, Diane (2014): *GDP: a brief but affectionate history*, Princeton: Princeton University Press.
- Crunchbase (2016): "Spotify". <https://www.crunchbase.com/organization/spotify>
- Dredge, Stuart (2015): "Apple Music interview: 'Algorithms can't do it alone – you need a human touch'", *The Guardian*, June 9.
- Economist (2009): "The triumph of the monthly bill", *The Economist*, October 8.
- Economist (2016): "The trouble with GDP", *The Economist*, April 30.
- Edwards, Jim (2016): "Spotify is causing a major problem for economists", *Business Insider*, September 26 <http://www.businessinsider.de/spotify-problem-for-economists-2016-9> (Accessed 15/03/17)
- Ernst, Wolfgang (2013): *Digital Media and the Archive*, Minneapolis: University of Minnesota Press.
- Felländer, Anna (2015): "The opportunities and challenges of digitalisation", *Digital Opportunities, Report from the Digitalisation Commission* (SOU 2015:65).
- Financial Times (2011): "Spotify bows to privacy pressure as Facebook sharing meets friction", blog post, September 29. <http://blogs.ft.com/tech-blog/2011/09/spotify-private-listening/> (Accessed 15/03/17)
- Fleischer, Rasmus (2015): "Towards a Postdigital Sensibility: How to get Moved by too Much Music", *Culture Unbound* 7, 255–269.
- Fleischer, Rasmus (forthcoming, 2017): "Nätutopier och nätdystopier. Om 2000-talets sökande efter internets väsen", Anders Burman & Lena Lennerhed (eds): *Samtider*, Göteborg: Daidalos.
- Galuszka, Patryk (2015): "Music Aggregators and Intermediation of the Digital Music Market", *International Journal of Communication* 9, 254–273.
- Garfinkel, Harold (1967): *Studies in Ethnomethodology*, Englewood Cliffs: Prentice Hall.
- Guardian* (2016): "Spotify in advanced talks to buy SoundCloud, reports say", 28 September. <https://www.theguardian.com/technology/2016/sep/28/spotify-soundcloud-acquisition-reports> (Accessed 15/03/17)
- Johansson, Sara (2015): "Så startade Spotify", *Centre for Business History*, February 2 <http://naringslivshistoria.se/bizstories-nyheter/naringslivshistoria/en-bra-ide-ara-borjan/> (Accessed 15/03/17)
- Kirschenbaum, Matthew (2008): *Mechanisms. New Media and the Forensic Imagination*, Cambridge, Mass: MIT Press.
- Marres, Noortje & Gerlitz, Carolin (2016): "Interface methods: renegotiating relations between digital social research, STS and sociology", *The Sociological Review* 64, 21-46.
- Morozov, Evgeny (2013): *To Save Everything, Click Here: The Folly of Technological Solutionism*, New York: Public Affairs.
- Music Week* (2009): "What does the future hold for Spotify?", *Music Week* 36, 15.
- Peoples, Glenn (2010): "Is Spotify really all that?", *Billboard*, 122(44), 20–21.
- Peoples, Glenn (2012): "Internet Radio: The Next Digital Underground", *Billboard*, July 20. <http://www.billboard.com/biz/articles/news/1084561/internet-radio-the-next-digital-underground> (Accessed 15/03/17)
- Pollack, Neal (2011): "The Celestial Jukebox", *Wired* 19(1), 74.

- Rogers, Richard (2013): *Digital Methods*, Cambridge, Mass.: MIT Press.
- Ruppert, Evelyn, Law, John & Savage, Mike (2013): "Reassembling social science methods: the challenge", *Theory, Culture & Society* 30(4), 22-46.
- Singleton, Micah (2016): "Does Spotify need to go after exclusive content?" February 18. <http://www.theverge.com/2016/2/18/11054460/spotify-exclusive-content-apple-tidal> (Accessed 15/03/17)
- Spotify (2011a): "Spotify introduces music to your social life", blog post, September 21. <https://news.spotify.com/se/2011/09/21/spotify-and-facebook/> (Accessed 15/03/17)
- Spotify (2011b): "What you share and how to control it", blog post, September 27. <https://news.spotify.com/se/2011/09/27/what-to-share/> (Accessed 15/03/17)
- Spotify (2011c): "What is Spotify?", archived version, July 14. <https://web.archive.org/web/20110714070521/http://www.spotify.com/int/about/what/> (Accessed 15/03/17)
- Spotify (2013a): Main US website, archived version, August 6. [https://web.archive.org/web/20130806170904/https://www.spotify.com/us/video-splash/?utm\\_source=spotify&utm\\_medium=web&utm\\_campaign=start](https://web.archive.org/web/20130806170904/https://www.spotify.com/us/video-splash/?utm_source=spotify&utm_medium=web&utm_campaign=start) (Accessed 15/03/17)
- Spotify (2013b): Main US website, archived version, May 2. <https://web.archive.org/web/20130502232221/https://www.spotify.com/us/> (Accessed 15/03/17)
- Spotify (2014): "Career Opportunities", archived version, January 22. <https://web.archive.org/web/20141221063949/https://www.spotify.com/se/jobs/opportunities> (Accessed 15/03/17)
- Sterne, Jonathan (2012): *MP3: The Making of a Format*, Durham: Duke University Press, 2012
- Vonderau, Patrick (forthcoming 2017): "Scaling Up: Streaming Services, High Finance, and Advertising Technology", Denise Mann (ed): *Content Wars*, New Brunswick: Rutgers University Press.
- Wade Morris, Jeremy (2015): *Selling Digital Music, Formatting Culture*, Oakland: University of California Press.
- Warren, Christina (2012): "Spotify Wants to Make Music Discovery Truly Social", Mashable, December 6. <http://mashable.com/2012/12/06/spotify-social-discovery/> (Accessed 15/03/17)
- Wikipedia (2017): "Spotify" <https://en.wikipedia.org/wiki/Spotify> (Accessed 15/03/17)
- Wikström, Patrik (2013): *The Music Industry*, Cambridge: Polity Press.
- Wikström, Patrik & DeFilippi, Robert (eds) (2016): *Business Innovation and Disruption in the Music Industry*, Cheltenham: Elgar.



# If the Song has No Price, is it Still a Commodity? Rethinking the Commodification of Digital Music

By Rasmus Fleischer

## Abstract

In music streaming services like Spotify, discrete pieces of music no longer has a price, as has traditionally been the case in music retailing, both analog and digital. This article discusses the theoretical and practical implications of this shift towards subscriptions, starting from a critical review of recent literature dealing with the commodification of music. The findings have a relevance that is not limited to music or digital media, but also apply more broadly on the study of commodification. At the theoretical level, the article compares two ways of defining the commodity, one structural (Marx), one situational (Appadurai, Kopytoff), arguing for the necessity of a theory that can distinguish commodities from all that which is not (yet) commodified. This is demonstrated by taking Spotify as a case, arguing that it does not sell millions of different commodities to its users, but only one: the subscription itself. This has broad economic and cultural implications, of which four are highlighted: (1) The user of Spotify has no economic incentive to limit music listening, because the price of a subscription is the same regardless of the quantity of music consumed. (2) For the same reason, Spotify as a company cannot raise its revenues by making existing customers consume more of the product, but only by raising the number of subscribers, or by raising the price of a subscription. (3) Within platforms like Spotify, it is not possible to use differential pricing of musical recordings, as has traditionally been the case in music retail. Accordingly, record companies or independent artists hence can no longer compete for listeners by offering their music at a discount. (4) Within the circuit of capital. Spotify may actually be better understood as a commodity producer than as a distributor, implying a less symbiotic relationship to the recorded music industry.

**Keywords:** capital, commodification, commodity-form, digital distribution, media industries, music, political economy, reification, Spotify, streaming, subscriptions.

Fleischer, Rasmus: "If the Song has no Price, is it Still a Commodity? Rethinking the Commodification of Digital Music", *Culture Unbound*, Volume 9, issue 2, 2017: 146–162. Published by Linköping University Electronic Press: <http://www.culture-unbound.ep.liu.se>

## Introduction

The concept of *commodification*<sup>1</sup> is widely used in the humanities and social sciences. Searching Google Scholar for the phrase “the commodification of” returns hundreds of titles asserting that culture as well as nature has been subject to the same process; the list of things commodified includes language, health, education, security, the body and the past. This article will review some recent literature on the commodification of music, but the findings have a broader relevance for attempts to study the political economy of digital media, and streaming services in particular. On a theoretical level, the aim is to sharpen the concept of commodification, which many scholars tend to use in a rather loose sense, sometimes synonymous with commercialization.

In more specific cases, the meaning of commodification might at first glance seem simple: something which is not a commodity then becomes one, typically by having a price tag attached to it. But this process tends not to leave the ‘thing’ untouched. In a longer historical perspective it stands clear that ‘the commodification of music’ does not simply indicate that a thing called ‘music’ is brought into the marketplace from the outside—the process also redefines the very meaning of ‘music’ (Fleischer 2012: 76-103).

In the tradition of critical theory, commodification is usually considered to be a “structural tendency in capitalism” (Jameson 2009: 257). A different approach is taken by cultural anthropologists who understand commodification as a “cultural and cognitive process” (Kopytoff 1986). In any case, speaking of commodification as a tendency or a process seems to necessitate a theory that can distinguish commodities from all that which is not (yet) commodified. Such a distinction, it may be noted, was pivotal to Karl Marx’ analysis of the capitalist mode of production. While he did begin *Capital* with a discussion of the commodity-form, his critique of political economy was conditioned on the observation that the classical economists were wrong in defining labour as a commodity.<sup>2</sup> But after him, critical theorists have usually been more keen to denounce commodification than to delineate it. This article will argue for the importance of having a precise concept of the commodity when trying to understand the political economy of digital media. Using the streaming music service Spotify as a case, my argument will be put forward in dialogue with some recent literature dealing with the commodification of music.

## Perspectives on the Commodification of Music

Ten years ago, ethnomusicologist Timothy D. Taylor remarked about the lack of theoretical treatments of how music is commodified. The commodity status of popular music, he wrote, “is so common that it seems to be self-evident, even

natural” (Taylor 2007, cf. Taylor 2006). Yet at the time of that observation, many observers regarded the practice of online file-sharing as subverting the commodification of music (Söderberg 2008: 8). The recorded music industry still betted on the success of digital retailers like Apple’s iTunes Store, selling music in the form of single-track downloads (IFPI 2007). But since then, the industry has been transformed anew by the rise of subscription-based streaming music services like Spotify.

This trajectory of music’s digitization has also inspired a new body of research, mostly within media studies, addressing how music is today put on the market. This literature can broadly be divided in two categories. On the one hand are the industry-oriented researchers who do not pretend to be critical, and whose research is chiefly based on interviews with music industry professionals. Their accounts tend to naturalize the commodity, asking not how music is commodified, but how the music commodity can be successfully “monetized” (Wikström 2009, Waelbroeck 2013, Johansson 2013, Galuszka 2015) On the other hand are those scholars who analyze the digitization of music through the lens of media history, zooming in on particular media formats, interfaces, affordances and social practices. Jonathan Sterne’s book *MP3* (2012) was pathbreaking in this respect; a similar approach has recently been taken by Jeremy Wade Morris in *Selling Digital Music* (2015) and also, to some degree, by the above-mentioned Timothy D. Taylor in *Music and Capitalism* (2015).

The latter three books share a similar approach to the concept of commodification. On the one hand, they all use the distinctly Marxian term *commodity-form*, though in a sporadic manner (Sterne 2012: 224, Morris 2015: 2, 213, Taylor 2015: 20–26) At the same time, all three books also rely to some degree on the influential theories of cultural anthropologists Arjun Appadurai and Igor Kopytoff, regarding how ‘things’ can move in and out of the commodity-state (Sterne 2012: 212–216, Morris 2015: 9–10, Taylor 2015: 10–11). These are two very different approaches, but none of the three authors provide any discussion about how to reconcile them – and in the end, none attempts to systematically apply any of the two theories on the turbulent reality of a digitizing music industry. The concept of commodification thus remains oddly undertheorized in the recent books by Sterne, Morris and Taylor. Taylor (2015, 21–24) does introduce a useful concept, “regimes of commodification”, identifying three in the history of commodified music:

We can consider music to exist in different regimes of commodification, all of which are still with us, though some are residual, some dominant, and some emergent: music as a published score, music as live sound at a public concert, and music as recorded sound in the form of player piano rolls or audio recordings in many other formats, analog or digital. (Taylor 2015: 21)

According to Taylor, all media technologies involving recorded sound belong to the same regime of commodification.<sup>3</sup> Morris, on the other hand, identifies a clear break between analog and digital: “The digital music commodity marks an evolution of the commodity-form” (Morris 2015: 11). What these two accounts have in common is that they rule out any substantial shift *within* the digital. Both authors are implicitly assuming that a download service like iTunes Store and a streaming service like Spotify are basically selling the same commodity, albeit in different wrappings.

This becomes especially problematic in the case of *Selling Digital Music*, as Morris claims to have written not a study of particular media but an inclusive analysis of “the digital music commodity”, and that his findings can be generalized as to apply for cultural commodities other than music. Throughout the whole book it is taken for granted that each file or song is a commodity. That makes sense when looking at a retailer like iTunes Store, where an apparent price tag is put on every single song. But pricing works very differently on streaming services like Spotify. In the latter case, an ordinary user will confront only one single price tag: 9.99 USD/EUR/GBP for a monthly subscription.

Only a few years ago, many economists and business journalists took for granted that internet would be ‘the great unbundler’: Instead of charging for a package, each item would be offered to consumers as a separate product. As a prime example of unbundling, these commenters pointed to Apple’s iTunes Store, selling singular tracks instead of albums (Akst 2005, Styvén 2007: 68, Carr 2008: 149–156, Elberse 2010, Pakman 2011). But this was soon followed by a massive re-bundling, as companies like Spotify began selling music as a subscription service. Instead of music being marketed as discrete units, each with one price tag, there is now only one product on offer: the subscription. As individual songs are made available within Spotify’s interface, without having any individual price, does it make sense to regard each song as one commodity? If not, in what sense is music commodified? Why does the theorization of the commodity matter for understanding the relationship between production and distribution, or the role of digital platforms, within contemporary capitalism?

These are the questions that I will attempt to answer in this article. Before entering a critical dialogue with Morris, I will take one step back for a brief reconstruction of the two abovementioned approaches to theorizing the commodity; on the one hand, the Marxian critique of political economy, on the other hand, the biographical approach of cultural anthropology. This will allow me to demonstrate how the theorization of the commodity matters, whether the focus of interest is music listening patterns, media business dynamics or the overall structure of digital capitalism.

## Critical Theory & the Commodity Form

‘Commodify’ is a relatively new verb. It was not established in English until the mid-1970s, probably as a translation of the Marxian expression “zur Ware werden” (Rose 2005, Haug 2010, Beech 2015: 231). Since then the term has made a rapid career in critical theory with Georg Lukács, Guy Debord and Karl Polanyi frequently invoked along Karl Marx as the classic theorists of commodification, though none of them ever used the term.

Marx did indeed dedicate the first chapter of *Capital* to an analysis of the commodity-form, describing it as the “economic cell-form” of the capitalist mode of production. Half a century later, in 1923, this line of analysis was dramatically amplified by George Lukács, declaring “the commodity-structure” to be the sole ground for “all the objective forms of bourgeois society together with all the subjective forms corresponding to them” (Lukács 1923/1971: 83). It is hard to overstate the significance of this intervention for the newly founded Frankfurt School and for what would be known as Western Marxism.

In this tradition of dialectical critique, the distinction between form and content is of highest importance. The commodity’s *content* might be a material object or an immaterial service, which gives it a particular use-value. But having a use-value is not enough. In order for something to take on the commodity-*form*, it must also have an exchange-value, i.e. it must stand in a market relation to other commodities. While the commodity is initially presented by Marx as the unity of use-value and exchange-value, this should not be mistaken for a final definition (a mistake made, for example, by Morris in *Selling Digital Music*). The Marxian critique of political economy is premised on categories – like commodity, value, labour and capital – whose interdependence can be properly deciphered only at the highest level of abstraction, as categories of a historically specific totality, commonly known as capitalism. Capitalism revolves around the abstraction of time, as commodity-producers compete with each other over minimizing the labour-time necessary in production (Postone 1993, Kurz 2012, Heinrich 2012). This process of abstraction is mediated by money. In chapter 3 of *Capital*, Marx (1867/1962: 109) declares money to be the “necessary form of appearance” for the abstracted labour-time. The fully developed commodity-form presupposes the existence of money. To put it very short, the Marxian definition of a commodity supposes that it has a price. On the other hand, it is not limited to tangible objects, but the commodity can as well be a service.

The analysis of the commodity-form does not end with the first chapter of *Capital*, as has often been assumed. It is not even completed in the first book. In book two of *Capital*, Marx proceeds from the economic cell-form to the circuit of capital. Here commodities and money appears as mere intermediary steps in an seemingly endless process of growth for growth’s sake. Capital for Marx is

not a thing, but can only exist in movement. The circuit presented by Marx has three stages and can be written out as a formula:  $M - C \dots P \dots C' - M'$ . In the first stage, money (M) is used to buy commodities (C): means of production and labour-power, together forming productive capital (P). The second stage consists of “productive consumption” in which the acquired commodities disappear in the creation of new commodities “of more value than that of the elements entering into its production”. In the third and final stage, the commodity-producer “returns to the market as a seller”, and capital is once again metamorphosing into money – the whole point being that there will now be more money than at the outset (Marx 1885/1963: 31). Every step in the process, however, implies a risk of capital’s devaluation. The value entering production will be conserved, and a surplus value added, only if the commodity producer succeeds in predicting the market correctly, so that the commodities can actually be sold at a profitable price (Manzerolle & Kjosen 2012).

At the center of this circuit – right in the middle of production – Marx finds a form of consumption. This “productive consumption” is not to be confused with individual consumption, which occurs outside the circuit of capital. Productive consumption means, for example, that raw materials are used up, that machinery and buildings are gradually worn down, and that the human time put into labour will never return again. But the capitalist mode of production, according to Marx, is organized so that value can be transferred from a preexistent commodity to a newly produced one. For something intangible (i.e. labour-power, care, knowledge, risk, or music) to become a commodity, it must first be transformed into an object. In critical theory, following Marx and Lukács, such objectification is known as reification. Thus, reification is a prerequisite for commodification, but does not in itself entail it. Fredric Jameson is consistent with Marx in concluding that it “is the market price which alone stamps an object as a commodity” (Jameson 2009: 257).

Proceeding from this, Jameson observes that in today’s academia, the analysis of commodification tends to diverge into two separate projects. On the one hand we find “abstract philosophical evocations” of “the organization of reality into things”; on the other, “specific empirical studies of the operations of markets in various fields” (Jameson 2009: 257–259). Bridging this theoretical gap is an important enterprise, to which Sterne, Morris and Taylor all contribute, to some degree. But this should not come at the expense of a blurred distinction between commodities and other objects.

If the commodity is defined by having a price, it should be fairly simple to tell whether a tangible object is commodified or not. But what about something like music? As a human practice, music can undoubtedly be subjected to different regimes of commodification, grounded in different modes of reification. This even



includes the possibility of commodification as resistance. Theodor W. Adorno saw modernist art as characterized by an attempt to turn commodification against itself (Martin 2007). If capitalism implies the tendential commodification of the entire social field, the modern artwork can, in Jameson's words, "only resist this external commodification by commodifying itself from the inside, by making itself over into a strange kind of mirror-commodity which is also an anti-commodity" (Jameson 2009: 264). Such an impulse might even be discerned in the "postdigital" tactics in some of today's musical subcultures; the commodification of material objects (like vinyl or cassettes) as an act of resistance against the attempts of platforms like Spotify to "commodify the void created by the lost materiality of music" (Fleischer 2015). In any case, the point of critical theory cannot be to just detect and denounce commodification in sweeping terms. One must always ask, what kind of object is made a commodity, before asking how it is done and how it can be possibly undone.

### **Cultural Anthropology and the Commodity Situation**

A different way of defining commodities can be found in the two contributions to the seminal volume *The Social Life of Things* by anthropologists Arjun Appadurai (1986) and Igor Kopytoff (1986). Their joint intervention aimed at discarding the structural analysis of the commodity-form in favour of the 'biographic' approach of "following the thing". The central idea is that an object, during the course of its 'social life', can be traced as it is "moving in and out of the commodity state" (Appadurai 1986). In other words, the commodity is not understood as a material thing, nor as a metaphysical form, but as a temporal phase in the life of a thing.

It cannot be denied that a strong "tangibility bias" is inherent in the approach of Appadurai and Kopytoff. Applying the biographical method to less tangible objects may cause confusion. In the current context, for example, it is far from certain how to interpret a music stream, i.e. a data sequence, as a 'thing'. Should the hardware device, the software interface and the sonic output then be regarded as three separate things, or as different aspects of the same thing? And while the metaphor of 'life' seems to imply birth and death, it remains unclear how to locate the beginning or the end of a thing's social life. Arguably, these uncertainties follow from the strong emphasis on exchange in this approach, as opposed to the Marxian emphasis on production (as productive consumption). David Graeber has also criticized Appadurai for "writing as if all exchanges are simply about *things* and have nothing to do with making, maintaining, or severing social relationships" (Graeber 2001: 32).

These problems notwithstanding, the essays by Appadurai and Kopytoff do provide clear criteria for deciding if a given object is in "the commodity state"

or not. Decisive is the kind of situation in which the object is currently found. According to Appadurai, “the commodity situation” is characterized by the need to sacrifice one object in order to access another one. In a modern economy, such exchange tends to involve money, although other forms of commodity exchange may also be found. As an alternative to the Marxian emphasis on production, Appadurai draws on Georg Simmel’s attempt to define “economic objects” based on their exchangeability:

Economic objects /.../ exist in the space between pure desire and immediate enjoyment, with some distance between them and the person who desires them, which is a distance that can be overcome. The distance is overcome in and through economic exchange, in which the value of objects is determined reciprocally. That is, one’s desire for an object is fulfilled by the sacrifice of some other object. (Appadurai 1986)

This emphasis on situation and context does not suffer from the abovementioned tangibility bias, but can be used as a benchmark to judge the “commodityness” of digital objects within a certain interface.

## Downloading, Streaming & Commodification

The preceding sections have presented two alternative theorizations of the commodity. These do not directly contradict each other, but approach the problem from very different angles. For the anthropologists, commodityness is all about the individual thing in a particular situation: if the thing can be enjoyed immediately, without the need to give up something else, it is simply not a commodity (though it may re-enter the commodity state at any time). From a Marxian perspective, on the other hand, the commodity must always be considered as part of a circuit, centred around the act of productive consumption; commodities are produced by way of other commodities, in order to be sold.

I will not attempt to reconcile these two approaches, but keep them both in mind as I proceed to revisit the notion of the music commodity. To put things in context, I will first provide a brief comparison of five modes of accessing music online (as of 2017), based on the assumption that you want to listen to one particular song. The simple question posed here will be if this song is delivered to you as a commodity.

1. If you choose to download the song from a digital retailer like iTunes Store, you will first have to pay its price. The price of each song is usually set at 0.99 USD/EUR/GBP, but may be higher or lower; a new release by a big artist may be priced higher due to high demand.

2. Alternatively, you can use a file-sharing network like Soulseek and download the song for free. The software for using this network does not expose the user for any advertising, and according to its developers, it is entirely financed by donations (Reitman 2016).

So far, the distinction seems clear: in the first case, the song is evidently a commodity, but not in the second. Matters however become a bit more complicated if you, instead of downloading a file to your computer, choose to stream the song directly from one of the many streaming services available.<sup>4</sup>

3. If you find the song on YouTube, you can listen to it for free, but it is quite likely that you will first be served an advertisement video. In this case, it is not the song itself that is sold to you, but your attention that is sold to advertisers. This kind of commodification is associated with the notion of the “audience commodity”, established by Dallas Smythe in the 1970s (Morris 2015: 99–100). At a metaphorical level, it does make sense to say that you “pay” for the song with your attention, or with the behavioural data that YouTube is collecting on you, but in strictly economic terms you are not paying for a commodity – rather you are the commodity being sold.

4. Just like in the previous case, Spotify Free offers streaming music at no charge but with interruptions for advertisement. Spotify does indeed differ from YouTube in several respects: it is centred around music, not video; it does not invite anybody to upload content; it is only available to registered users. None of these differences seem to alter the basic mechanism of commodification. But a closer examination may indicate that advertisement fulfils partly different functions in the two services. Advertisers are indeed paying Spotify have a message delivered to an audience, but in Spotify’s current “freemium” business model, advertisement revenues are arguably less important than the annoyance that users experience when the music is interrupted by ads. The main function of the ads is to prompt users to pay for a subscription.

5. In order to listen to music without interruptions, you can subscribe to Spotify Premium for a monthly 9.99 USD/EUR/GBP. In fact, this price is the only price that will ever confront an ordinary user of Spotify. The one and only commodity sold to consumers by Spotify is the subscription. The song itself is simply not delivered as a commodity to users of Spotify. It does not exist in what the cultural anthropologists would call a commodity-situation. As it does not have a price, it also does not fulfil the Marxian definition of a commodity.

## The File-Centric Understanding of the Digital Music Commodity

Jeremy Wade Morris' recent book *Selling Digital Music* (2015) is in many ways an excellent tracing of "the music commodity's re-tuning" in the transition from "compact discs to music as individual digital files" (Morris 2015: 2, 198–199). Settling there, however, it consolidates a particular kind of atomism, what I would call a file-centric understanding of the digital music commodity. This is problematic, since the author claims a broad validity of his findings, not only for every digital music service, but for "the process of cultural commodification" in general (Morris 2015: 5). "The digital music commodity is an object in its own right", and "a discrete sonic entity", Morris postulates at the outset. (Morris 2015: 3, 16) Throughout the book it is taken for granted that this commodity is essentially a file, representing a song (Morris 2015: 11, 192, 209–210) Sometimes the digital music commodity is equated with the simple audio data, other times it is presented as a bundle of audio data and metadata. What is never questioned is the notion of the song as musical atom being the object of commodification.

The interface, consisting of hardware and software, is what makes this object visible and audible, Morris writes: "It is the point where user and commodity meet." (Morris 2015: 18) In other words, the interface is not analysed as a commodity, nor as a commodifying device, but rather appears as an arena where the already commodified objects appear. This raises a substantial question about the definition of a commodity. Though he employs the Marxian notion of the commodity as *form*, Morris distances himself from Marx by stating that price "is not its distinguishing feature" (Morris 2015: 2, 8, 11, 20, 192, 210–213). And when looking for "the very core of the commodity", he finds not form but matter: a data file in a particular format (Morris 2015: 195). This view seems to be more in line with the anthropology of material culture, and early in the book Morris does indeed evoke Appadurai and Kopytoff to define commodities "as artifacts in a particular situation, the commodity situation" (Morris 2015: 9–10). But in his subsequent analysis of digital interfaces, Morris abstains from asking whether such a situation is at hand. At certain points, he follows a wholly different path, defining the commodity neither as form nor as situation, but by reference to a subjective "sense of ownership" (Morris 2015: 20). This theoretical confusion results in an inability to ever state that something is not a commodity. Ultimately, *Selling Digital Music* fails at its explicit aim: telling the "story of how the commodity form changes through digitization and why this matters for the music and media we love" (Morris 2015: 2).

Morris seems to assume that it is the same commodity which is made available by "iTunes, Spotify, and many other digital retailers" (Morris 2015: 210). The formulation reveals a further assumption, namely that Spotify is essentially a retailer, i.e. a distributor of commodities that has already been produced el-

sewhere. Retailing implies buying commodities in order to resell (or rent out) the same commodities. If a company is selling commodities different from those it has bought, it is evidently involved in some kind of production. But the distinction between production and distribution does not appear in *Selling Digital Music*, due to the lack of a clearly defined notion of the commodity. Whether a streaming service's activities are regarded as distribution or production does however have analytical repercussions.

In this section my argument has been developed as a critique of some implicit assumptions in *Selling Digital Music*. Now it is time to turn towards self-critique. After all, this article comes out of a research project about Spotify with the outspoken aim of 'following files in digital music distribution'. But what if, as already indicated, Spotify is better understood not as a music distributor but as a producer? Then a new question will arise: what commodity is Spotify producing?

## Producing the Branded Musical Experience

Spotify began as a streaming service based on the on-demand doctrine, tailored for individuals who already knew exactly which pieces of music they wanted to hear: "Millions of tracks, any time you like [...] Just help yourself to whatever you want, whenever you want it." This was soon to change (Fleischer & Snickars 2017). In 2013 the company adopted a new slogan: 'The right music for every moment'. The current aim is to provide a programmed soundtrack to each 'moment' of the user's day. Asked to explain the core of his business, Spotify's CEO Daniel Ek explains: "We're not in the music space—we're in the moment space" (Seabrook 2014). This is part of a wider shift that has been brilliantly analyzed by Jeremy Wade Morris in a recent article co-authored with Devon Powers. Spotify and other "outlets that once primarily concerned themselves with distribution are now increasingly in the business of promotion, curation, user experience and analytics", they write. These services want to sell "music as an affective experience rather than as individual songs" (Morris & Powers 2015).

Indeed, streaming services are not the first to commodify music by embedding it in a broader 'experience'. One need only to think about music festivals, MTV, or Apple's marketing of the iPod. Still, the mechanism of commodification does differ between digital music services. Apple's iTunes uses the experience to sell commodified pieces of music, each piece being a commodity with a price. Spotify uses the music to sell a commodified experience, bundled together as one commodity. As already stated, the main problem with *Selling Digital Music* is that Morris here postulates a uniform logic of "the digital music commodity". The article co-written by Morris and Powers, on the contrary, points towards a transformed commodity-structure *within* the digital.

Ultimately, we suggest that digital music services no longer sell discrete musical objects, nor do they focus exclusively on content offerings. Instead, services sell branded musical experiences, inviting consumers to see themselves and their attitudes, habits and sentiments about music reflected by the service they choose to adopt. (Morris & Powers 2015)

It can be noted, however, that the concept of ‘the music commodity’, so prominent in *Selling Digital Music*, is all but absent in the recent article by Morris and Powers. It deserves to be re-introduced in the context of branded musical experiences, so that the transformed commodity-structure of digital music can be properly situated within the circuit of capital.

## Conclusions

Music may be commodified in several ways: as published score, as live performance, as recorded sound. These are the three “regimes of commodification” mentioned by Timothy D. Taylor. However, the argument can be made that streaming music services like Spotify represent yet another regime. This fourth regime of commodification – music as part of a branded experience – is not new, but it may be argued that it is now becoming dominant. When music is not sold as individual pieces or events but is embedded in a personalized service, the distinction between production and distribution tends to blur. To enable a critical analysis of commodification it is necessary to first ask what commodity that is actually being sold to consumers, in this case by Spotify.

In this article, I have criticized Jeremy Wade Morris’ concept of ‘the digital music commodity’ because it assumes that each song, whether it is downloaded or streamed, is distributed as a commodity. In the case of download retailers like iTunes Store, that is correct. But the user of Spotify will never see a price tag on a song, and never need to give up anything in order to consume one more song. Accordingly, the user of Spotify does not consume each song as a commodity. That is the only possible conclusion, whether one prefers to define the commodity structurally (Marx) or situationally (Appadurai, Kopytoff).

Spotify is indeed not selling discrete pieces of recorded sound, neither by the song, by the album, or in any other portioning. It is not putting millions of commodities on the market, but offering consumers only one commodity: the subscription. This product is a bundle that includes not only access to all songs in the catalog but also the maintenance of a personalized profile connected to a variety of playlists tailored for pre-defined activities, helping the user to navigate through an abundance of music without first having to choose which songs to play. This

marks a clear difference to music retailers, analog or digital, which simply sell pieces of recorded music. While these retailers are indeed in the business of music distribution, streaming services are moving away from the distributive function, attempting to occupy another place in the circuit of capital. Currently it may make more sense to regard Spotify as a commodity producer.

According to Karl Marx, commodity production includes the "productive consumption" of labour-power and means of production. In order to produce the "branded musical experience", Spotify has to acquire on the one hand labour-power (from network engineers to graphic designers and music curators), and on the other hand, means of production (including hardware, bandwidth, and music licenses). Of all these expenses, music licenses make up the by far largest part. The productive consumption of licenses does also put a limit to the possible economies of scale. For every new user, Spotify has to pay money to the rights holders, though the details of the licensing deals are trade secrets, unavailable for independent researchers.

When regarded as a commodity producer, Spotify seems to have more in common with traditional broadcasters than with any sort of music retailer. The radio listener also does not receive commodities over the air when music is played – rather, music is used to harness the attention of the listener in the production of the audience commodity that is sold to advertisers. For this, the radio broadcaster has to buy a music license from a copyright collecting society, and the pricing of this license does not discriminate between different pieces of music; typically, a flat fee is paid by the minute. Economically, this is very different from the record shop (or the download shop), where a new and popular recording may cost more than an old one. If radio broadcasters and record shops represented two different regimes of commodification in the 20th century, Spotify's business model may be better understood as a mutation of the traditional broadcasting model than as a new form of retail.

It can hardly be denied that Spotify is a heavily commodified environment. But compared to the selling of individual songs by retailers like iTunes Store, streaming services like Spotify have a very different way of commodifying music. The de-commodification of individual recordings (at the consumer's side), now coincides with the re-commodification of music as an experience. This shift in music's commodity structure does indeed matter for the individual music listener, just like it matters for a company like Spotify, and for the recorded music industry at large. Furthermore, it has repercussions for the broader analysis of how to situate digital platforms in the circuit of capital. I will sum up this article by briefly motivating these four points.

1. The user of Spotify does not confront songs or albums as commodities. The only commodity the user is invited to consume is the subscription itself. Everybo-

dy pays the same monthly fee to access the same service. When listening to music on Spotify as a subscriber, there is no longer a commodity-situation at hand, hence no incentive to limit music listening, neither as in listening time, nor as in number of songs accessed.

2. This commodity-structure also determines the possibility of growth for a company like Spotify. There is no real point for it to make existing subscribers consume more music, as these will pay the same amount per month regardless if they listen to 10 or 10,000 songs. The two remaining ways to increase revenue would then be to raise the cost of a subscription, or to raise the number of subscribers.

3. For the recorded music industry, this creates different incentives as compared to the sale of music as discrete units. Most importantly, it removes the possibility for differential pricing. As long as discrete units of music are sold to consumers via a distributor, a new or exclusive recording can be made more expensive, while an upcoming artist can choose to offer music much cheaper, or at no cost at all, thereby gaining a better chance to compete. Even a digital retailer like iTunes Store allows for a limited degree of differential pricing, but Spotify does not. Price competition has now been effectively abolished. The rights holder cannot try out different pricing strategies, but only choose whether to be on the platform or be absent.<sup>5</sup>

4. It has usually been taken for granted that the record industry produces a commodity, which is then distributed by services like Spotify. This could imply a symbiotic relationship based on a shared interest in keeping up the value of recorded music. Matters appear a bit different if Spotify is instead considered a producer of a new commodity, the branded musical experience. Then music (commodified as licenses) is simply one of several inputs, albeit the most important one, to the production process. Accordingly, the devaluation of recorded music would be in the interest of Spotify. If it is true that Spotify is 'not in the music space' but 'in the moment space', this would also mean that Spotify is not just competing with other music streaming or downloading services, but potentially with a broader range of services promising to personalize and optimize everyday activities like studying, exercising or partying.

Music is still commodified by Spotify. But as a commodity, 'music' can mean very different things. It is a concept too vague for allow for a precise analysis of its political economy. Spotify buys music in the form of copyright licenses, bundling it, adding new features, and then sells music as a personalized experience. The market for such a service can be expected to work differently from the market for discrete pieces of recorded sound. Thus, simply asserting that "music" is commodified does not say much at all. When analyzing commodification it is always necessary to ask what kind of object that is made a commodity. A good start is to check where the price tag is hanging.



**Rasmus Fleischer** is a postdoctoral researcher based at the Department of Economic History, Stockholm University. His research interests are located in the intersections between culture and economy, as well as technology and politics. Most of all, he has explored 20th century media history, transformations of copyright and the commodification of music. E-mail: [rasmus.fleischer@ekohist.su.se](mailto:rasmus.fleischer@ekohist.su.se)

## Notes

<sup>1</sup> *Commodification* is not the same as *commoditization*. While the first term has to do with a movement into the market, the latter stands for a change in the mechanisms of competition on the market. Commoditization implies that buyers no longer pay attention to the particular brand of a product, but regards it more as an undifferentiated material like sugar, oil or electricity (the alternative sense of "commodity"). This concept was established in business theory in the 1990s. Before that, the two terms were used interchangeably.

<sup>2</sup> According to Marx, it is not *labour* but *labour-power* that is commodified under capitalism. Indeed, the whole riddle of exploitation is located in this distinction. Labour-power is the capacity to perform labour. By buying labour-power, the capitalist can exploit labour and retain the results of surplus-labour (Marx 1867/1962: 181-191, Clarke 1982: 89, Postone 1991: 270, Heinrich 2012: 91).

<sup>3</sup> It can be noted that Taylor (2015) departs from his own earlier claim (Taylor 2006), that digital media was bringing about a wholly new regime of commodification, with some new openings toward a de-commodified experience of music – a line of thought reminiscent of Jacques Attali (1985, 133–148). Taylor's distinction between the residual, the dormant and the emergent is borrowed from Raymond Williams (1973).

<sup>4</sup> Technically, streaming *is* downloading. The only difference is what happens to the data after it is downloaded: either it is placed in volatile memory for immediate playback (streaming), or it is saved in a more permanent way (downloading).

<sup>5</sup> In a strict economic sense, it could be argued that price competition has not been abolished, but only reduced to the binary choice between charging something and charging nothing. Those who want to promote their music by charging nothing for it may choose distribution platforms like YouTube or SoundCloud.

## References

- Akst, Daniel (2005): "Unbundles of Joy", *The New York Times*, Dec. 11.
- Appadurai, Arjun (1986): "Introduction: Commodities and the Politics of Value", Arjun Appadurai (ed.), *The Social Life of Things: Commodities in Cultural Perspective*, Cambridge: Cambridge University Press, 3–64.
- Attali, Jacques (1985): *Noise. The Political Economy of Music*, Minneapolis: University of Minnesota Press.
- Beech, Dave (2015): *Art and Value. Art's Economic Exceptionalism in Classical, Neoclassical and Marxist Economics*, Chicago: Haymarket Books.
- Carr, Nicholas (2008): *The Big Switch: Rewiring the World, from Edison to Google*, New York: Norton.
- Clarke, Simon (1982): *Marx, Marginalism and Modern Sociology: from Adam Smith to Max Weber*, London: Macmillan.
- Elberse, Anita (2010): "Bye Bye Bundles: The Unbundling of Music in Digital Channels", *Journal of Marketing* 74:3.
- Fleischer, Rasmus (2012): *Musikens politiska ekonomi. Lagstiftningen, ljudmedierna och försvaret av den levande musiken, 1925–2000*, Stockholm: Ink bokförlag.
- Fleischer, Rasmus (2015): "Towards a Postdigital Sensibility: How to get Moved by too Much Music", *Culture Unbound* 7, 255–269.
- Fleischer, Rasmus & Pelle Snickars (2017): "Discovering Spotify – A Thematic Introduction", *Culture Unbound*, 9:1, 1–16.
- Galuszka, Patryk (2015): "Music Aggregators and Intermediation of the Digital Music Market", *International Journal of Communication* 9.
- Graeber, David (2001): *Toward an Anthropological Theory of Value. The False Coin of our own dreams*, New York: Palgrave.
- Haug, Wolfgang Fritz (2010): "Kommodifizierung", in Wolfgang Fritz Haug, Frigga Haug and Peter Jehle (eds.), *Historisch-kritisches Wörterbuch des Marxismus, Bd 7/2, Knechtschaft bis Krise des Marxismus*, columns 1243–1255, Hamburg: Argument.
- Heinrich, Michael (2012): *An Introduction to the Three Volumes of Karl Marx's Capital*, New York: Monthly Review Press.
- IFPI (2007): "Digital Music Report 2007", available at: <http://www.ifpi.org/content/library/digital-music-report-2007.pdf> (accessed 09/09/16).
- Jameson, Fredric (2009): *Valences of the Dialectic*, London: Verso.
- Johansson, Daniel (2013): "From Products to Consumption. Changes on the Swedish Music Market as a Result of Streaming Technologies", Working Paper, Linnaeus University, Sweden.
- Kopytoff, Igor (1986): "The Cultural Biography of Things: Commodization as Process", Arjun Appadurai (ed): *The Social Life of Things. Commodities in a Cultural Perspective*, Cambridge: Cambridge University Press, 64–91.
- Kurz, Robert (2012) *Geld ohne Wert: Grundrisse zu einer Transformation der Kritik der politischen ökonomie*, Berlin: Horlemann.
- Lukács, Georg (1923/1971): "Reification and the Consciousness of the Proletariat", in *History and Class Consciousness: Studies in Marxist Dialectics*, London: Merlin Press.
- Martin, Stewart (2007): "The Absolute Artwork Meets the Absolute Commodity", *Radical Philosophy* 146, 15–25.
- Marx, Karl (1867/1962): *Das Kapital: Kritik der politischen Ökonomie. Erster Band: Der Produktionsprozeß des Kapitals*, reprinted as Marx-Engels-Werke (MEW), volume 23, Berlin: Dietz Verlag.
- Marx, Karl (1885/1963): *Das Kapital: Kritik der politischen Ökonomie. Zweiter Band: Der Zirkulationsprozeß des Kapitals*, reprinted as Marx-Engels-Werke (MEW), volume 24, Berlin: Dietz Verlag.

- Manzerolle, Vincent R. & Atle Mikkola Kjösen (2012): “The Communication of Capital: Digital Media and the Logic of Acceleration”, *Triple C* 10 (2), 214-229.
- Morris, Jeremy Wade (2015): *Selling Digital Music, Formatting Culture*, Oakland: University of California Press.
- Morris, Jeremy Wade & Devon Powers (2015): “Control, Curation and Musical Experience in Streaming Music Services”, *Creative Industries Journal* 8 (2), pp. 106–122.
- Pakman, David (2011): “The Unbundling of Media”, blog post available at: <http://www.pakman.com/2011/04/15/the-unbundling-of-media/> (accessed 16/03/17).
- Postone, Moishe (1993) *Time, Labor, and Social Domination. A reinterpretation of Marx's Critical Theory*, Cambridge: Cambridge University Press.
- Reitman, Rainey (2016) “The Kafkaesque Battle of Soulseek and PayPal, and Why Free Speech Defenders Should be Worried About Payment Networks”, *Electronic Frontier Foundation*, February 25, available at: <http://www.ifpi.org/content/library/digital-music-report-2007.pdf> (accessed 16/03/17).
- Rose, Carol (2005): “Afterword: Wither Commodification”, Martha M. Ertman & Joan C. Williams (eds.): *Rethinking Commodification: Cases and Readings in Law and Culture*, New York: New York University Press, 402–428.
- Seabrook, John (2014): “Revenue Streams. Is Spotify the Music Industry’s Friend or its Foe?”, *The New Yorker*, November 24, 2014.
- Sterne, Jonathan (2012): *MP3. The Meaning of a Format*, Durham: Duke University Press.
- Styvén, Maria (2007): *Exploring the Online Music Market. Consumer Characteristics and Value Perceptions*, diss. Luleå tekniska universitet.
- Söderberg, Johan (2008): *Hacking Capitalism. The Free and Open Source Software Movement*, New York: Routledge.
- Taylor, Timothy D. (2006): “Music + Digital Culture: New Forms of Consumption and Commodification”, Paul Messaris & Lee Humphreys (eds.): *Digital Media: Transformations in Human Communication*, New York: Lang, 87–96.
- Taylor, Timothy D. (2007): “The Commodification of Music at the Dawn of the Era of ‘Mechanical Music’”, *Ethnomusicology* 51 (2), 281–305.
- Taylor, Timothy D. (2015): *Music and Capitalism. A History of the Present*, Chicago: The University of Chicago Press.
- Waelbroeck, Patrick (2013): “Digital Music: Economic Perspectives”. Available at SSRN: <https://ssrn.com/abstract=2249690> or <http://dx.doi.org/10.2139/ssrn.2249690>
- Wikström, Patrik (2009): *The Music Industry: Music in the Cloud*, Cambridge: Polity.
- Williams, Raymond (1973) “Base and Superstructure in Marxist Cultural Theory”, *New Left Review* I (82).



# Tracking Gendered Streams<sup>i</sup>

By Maria Eriksson & Anna Johansson

## Abstract

One of the most prominent features of digital music services is the provision of personalized music recommendations that come about through the profiling of users and audiences. Based on a range of “bot experiments,” this article investigates if, and how, gendered patterns in music recommendations are provided by the streaming service Spotify. While our experiments did not give any strong indications that Spotify assigns different taste profiles to male and female users, the study showed that male artists were highly overrepresented in Spotify’s music recommendations; an issue which we argue prompts users to cite hegemonic masculine norms within the music industries. Although the results should be approached as historically and contextually contingent, we argue that they point to how gender and gendered tastes may be constituted through the interplay between users and algorithmic knowledge-making processes, and how digital content delivery may maintain and challenge gender relations and gendered power differentials within the music industries. Seen through the lens of critical research on software, music and gender performativity, the experiments thus provide insights into how gender is shaped and attributed meaning as it materializes in contemporary music streams.

**Keywords:** gender, music, algorithms, streaming, digital methods

Eriksson, Maria & Anna Johansson: “Tracking Gendered Streams”, *Culture Unbound*, Volume 9, issue 2, 2017: 163–183. Published by Linköping University Electronic Press: <http://www.cultureunbound.ep.liu.se>

## Introduction

Being a user of contemporary online services usually means being “profiled.” In broad strokes, the profiling of online users—what is commonly described as a “personalization” of data flows—involves the recording of user behaviors, the approximation of user tastes and preferences, and the delivery of content assumed to fit an individual’s needs and wishes. As Mike Featherstone (2006: 592) has argued, similar treatments of data have historically made people “identified and individuated...as whole populations, their bodies and life histories became documented, differentiated, and recorded.” Rooted in the practices of nation states, the profiling of populations is increasingly performed by digital media corporations through the treatment of online data. The technical systems that generate this kind of profiling may be understood as “infrastructures of taste formation” with a profound capacity to shape cultural encounters in both descriptive and prescriptive ways (Morris 2012; Beer 2013: 97). Content deliveries that are based on user profiling indicate how the preferences of users are imagined and they also guide the choices users can make. Thus, they may also serve to reproduce certain behaviors.

Working on a project concerned with streaming music, we wanted to explore how the streaming service Spotify may be constitutive of user identities and taste constellations related to gender. Spotify is currently one of the largest providers of digital music, and a company that hosts the music consumption of over 100 million users around the world. Due to its popularity, Spotify is a great example of a new media service that is “central in shaping our everyday lives and in ordering our routine experiences” (Beer 2013: 1). In a previous study, for example, we explored how Spotify evokes normative temporalities, neoliberal subjectivities, and functional approaches to music through its ways of greeting users (Eriksson and Johansson, forthcoming). In this article, our main interest instead lies in the way gender comes to matter in Spotify’s music recommendations.

We argue that to fully understand the cultural significance of contemporary online media, scholars in the humanities need to directly engage with the digital systems (algorithms, software assemblages, data flows) that organize media usage. In order to explore whether and how gender is produced through Spotify, we therefore set up a case study that investigated 1) the extent to which users’ gender identification had an impact on the provided music recommendations, and 2) the gender ratio of artists recommended to Spotify users. Seen through the lens of previous research on software, music and gender, this provides an understanding of whether algorithmic user profiling and music recommendations are constitutive of gender and gendered tastes, and of how digital content delivery can maintain and challenge gender relations and gendered power differentials within the music industries.

## Algorithms and User Profiling

On Spotify, algorithmic processes are heavily involved in framing, moving, assembling, and contextualizing music in ways that affect who sees it, and how it is perceived. Over the past few years, studies of online recommendation systems and algorithmic knowledge production have proliferated (Introna & Nissenbaum 2000; Mackenzie 2006; Amoores & Piotukh 2016; Kitchin 2017), together with a broader interest in the politics and cultural implications of code and software (Fuller 2003; Berry 2011; Manovich 2013). Algorithms have received attention as “a kind of invisible structural force that plays through into everyday life” (Beer 2013: 69), for instance by providing personalized recommendations of online content. Constituting a new form of power, it has been suggested that the design of algorithms and algorithmic procedures foster certain cultures, ideologies and identities. Mager (2012), for example, has demonstrated how search engines embody an “algorithmic ideology” closely connected to capitalist modes of production, while others have shown how Google perpetuates stereotypes related to race and gender through its deliveries of online content (Olofsson 2015; Baker & Potts 2013). Because algorithms serve as cultural intermediaries that help content “find us”, they also affect how cultural capital is acquired and how taste is shaped (Beer 2013; see also Morris 2015). In this way, algorithms act alongside other cultural intermediaries like record labels, music aggregators, music critics and retailers within the music industries (cf. Drew 2005; Galuszka 2015). Together, such actors perform the task of mediating and creating meaning around music and its audiences. Thus, algorithms are not the only means by which music acquires meaning and fans become profiled, but their involvement in musical processes require further attention.

Currently, critical research on algorithmic filtering systems vary with respect to the extent that algorithms are attributed agency in and of themselves – ranging from studies that theorize algorithms in nonanthropocentric terms; that is, as more-than-human actors which actively shape social life (Parisi 2013; Dixon-Roman 2016), to studies focusing on how algorithms are developed, deployed and attributed meaning in specific cultural contexts (cf. Seaver 2013; Gillespie 2014). While we take inspiration from the former strand’s emphasis on the tangible effects and manifestations of algorithms, we primarily approach algorithmic systems as socio-technical configurations which are inscribed with, and hence performative of, particular world-views when activated by users (cf. Kitchin 2017).

As discussed by Cheney-Lippold (2017), algorithmic ways of organizing content have implications for the construction of gender as well as other social categories. Because of their sorting mechanisms, algorithms both infer and define the meaning of such categories in a process which could be likened to a form of soft biopolitics; that is, a mode of governing bodies and populations. By “deciding

what it is that the individual encounters, and what it is that they are making a decision about” (Beer 2013: 96), content recommendation algorithms work as prescriptive entities that simultaneously predict user preferences and contribute to the shaping of practices and identities—including gendered forms of existence. Much of this prescriptive work takes place through the profiling of users. As Amanda Modell (2015: 9) has put it, the code and algorithms that power online music deliveries create “positional relations between bodies and sets of music from a seemingly objective standpoint.” In this way, “code mediates technoscience and consumer self-knowledge” (ibid.: 4). Therefore, it matters how a service like Spotify imagines and responds to its users.

### *Software and gender*

Placing our study in the context of current research in the emerging fields of software studies and new media studies, we take the stance that technology is inherently social and political. Feminist studies of technology have emphasized how technology is shaped by the circumstances in which it occurs, and, hence, how “gender relations can be thought of as materialized in technology, and gendered identities and discourses as produced simultaneously with technologies” (Wajcman 2007: 293; for further discussion see also Haraway 1997; Hayles 1999; Sundén 2015). Software, algorithms, and data, we argue, are sociotechnical configurations with discursive as well as material components that are situated in gendered contexts and hence embedded in particular values and gender discourses. According to Bivens and Haimson (2016: 1), design decisions “determine where—in the multiple layers of software—gender appears as a category, how it is materialized within code and activated within software processes, and for what purposes it is deployed.” This suggests that digital technologies and infrastructures may be complicit in the reproduction of hegemonic gender relations, but also that they can be put to use for unintended purposes and with unanticipated consequences.

Our study is further informed by the notion of gender performativity (Butler 1990, 2004), which we argue can be fruitfully combined with an understanding of algorithmic systems as performative entities. Following Butler, we take gender to be the material effect of regulatory discourses, and something which only exists to the extent that it is repeatedly enacted in social practice. Binary gender—naturalized notions of masculinity and femininity—are thus seen as a contingent set of positions which are iteratively produced and reproduced through “a stylized repetition of acts” that builds on prescriptive conventions (Butler 1990: 179). By citing existing norms, these can “be exposed as non-natural and nonnecessary when they take place in a context and through a form of embodying that defies normative expectations” (Butler 2004: 218).<sup>1</sup>

While Butler's notion of gender performativity mainly focuses on how gender is brought into existence by linguistic and embodied practices, she claims that performative power is also exercised through "organizations of human and non-human networks, including technology" (Butler 2010: 150). In the specific context of gender and software, a similar line of thinking has also inspired studies of how programmed configurations of gender materialize on social media services, and how gendered technologies are mutually shaped by users and software designers (Bivens & Haimson 2016). In a study of design decisions related to gender on Facebook, for example, Bivens (2015: 2) shows how code and software can be seen as producing "the conditions for gendered existence" by normalizing a binary logic. Algorithmic music recommendations, we suggest, represent another potentially illustrative example of how software comes to regulate social life.

## **Gender and music**

Importantly, algorithmic music recommendations are tightly interwoven with the music industries at large, where gender is of major significance. The persistent male domination in the music industries has been noted by several scholars over the years, hence acknowledging the marginalization of women in music production and the ways in which gender conventions and ideologies affect music practices (Frith & McRobbie 1991; Whiteley 2013; Gavanoas & Reitsamer 2013). For instance, it has been noted how recording studios, tour buses and guitar shops are constructed as masculine contexts (Bayton 2013; Leonard 2007) and how technological mastery is strongly associated with male expertise (Gavanoas & Reitsamer 2013). As a consequence, female musicians are often designated as exceptions, thus normalizing the male performer (Leonard 2007; Gadir 2016).

Prescriptive ideologies of masculinity and femininity are also "bound up with particular musical styles" (Whiteley 2013: xix), and gender ideologies inform the valuation of different music genres for performers and fans alike. Here, the rock/pop binary is arguably the most well-cited: scholars have problematized the ways in which pop is typically attributed feminine characteristics and a mainstream, commercial orientation, whereas rock music is seen as masculine, authentic and of higher value—thus reproducing the marginalization of women in rock (Railton 2001). Similarly, gender relations are played out and naturalized in relation to other genres, manifested for example in the co-construction of white masculinity and indie rock (Bannister 2006), the devaluation of women and gay rappers (Jeffries 2011; Berggren 2013), and male dominance in DJ culture (Gadir 2016). Typically, such studies point to the interconnectedness of gender, race, class and sexuality in relation to different music styles.

At the same time, feminist scholars have noted how patterns of domination



and exclusion in popular music are negotiated and subverted, and thus how music can also act as a vehicle for transformation of gendered and sexual subjectivities (Pough 2004; Whiteley & Rycenga 2013). It has been suggested that digital technology, especially social media, enable women's self-production of music (Choi 2016), but also that digitalization perpetuates the view of technologies as domains of masculinity and male expertise, in the context of music production (Gavanas & Reitsamer 2013) as well as consumption (Werner & Johansson 2016). However, while a number of studies have explored the impact of emerging streaming technologies on music distribution and music practices (Morris 2015; Kjus 2016; Maasø 2017), few have specifically addressed the significance of gender in these processes. Exceptions are for example Werner and Johansson (2016, 178), who discuss how "music and technology emerge as gendered in talk about contemporary online music use." The present study contributes to the field by focusing on the embeddedness of gender discourses in media technologies themselves, and by suggesting innovative methods for the study of gender, music and digital technology.

## **Bot Methods**

For the purpose of this study, we arranged an experiment that explored similarities and differences in music recommended to Spotify users registered as male and female.<sup>2</sup> The experiment was carried out with the help of system developers Roger Mähler and Johan Von Boer at Humlab, Umeå University. Grounded in digital methods and the use of software affordances as research tools (Rogers 2013), the experiment involved the creation of programmed informants—essentially coded scripts, or bots—that were instructed to behave like ordinary users.<sup>3</sup> In taking such an approach, we wanted to directly engage with Spotify's software, rather than studying its dynamics as they are mediated by its company representatives in traditional industry interviews or public documents. In alignment with Evelyn Ruppert, John Law and Mike Savage (2013), we argue that scholars who research digital technologies need to get their hands dirty and explore the affordances of digital devices and how they collect, store, transmit, sample, and forge social relations. By experimentally engaging with digital technologies and testing the boundaries of what can be known about Spotify's recommendation systems from the outside, we hope to contribute to such an emerging conversation.

Currently, little is known about how Spotify's recommendation algorithms operate in relation to gender, although clues might be drawn from blog posts such as "Gender Specific Listening", written in 2014 by Paul Lamere, director of the developer platform for the Spotify-owned company The Echo Nest. The Echo Nest has helped manage Spotify's music recommendations since many years and in his

text, Lamere argues that gender is one of the key demographic variables that can say something about a user's taste in music. Based on the analysis of historical user data, Lamere explains that identifying and eliminating artists that are "gender skewed" is one of the prime strategies by which gender might be used to modify music recommendations. His argument reveals that notions of "gender specific music tastes" exist among software developers, but the text does not confirm that gender-adapted recommendation schemes are actually put to use in the Spotify service.<sup>4</sup> This is what we set out to investigate in our research.

In the experiment, we began by first registering 288 new Spotify accounts.<sup>5</sup> These accounts had the exact same settings (address, date of birth, home address, privacy settings etc), but half of the users were registered as male and half as female. At the time, male/female were the only gender options available upon registration to the service (In September 2016, Spotify began to roll out the possibility of registering a third gender category in select countries – an issue we will return to in our discussion.)<sup>6</sup> The 288 accounts were then divided into four groups, and the bots in each group were instructed to listen to music from one of the genres rock, gospel, rnb/hiphop, and dance/electronic music. The genres were borrowed from Billboard's global hit lists at the time of the study, and the users were instructed to listen to the ten most popular songs on Billboard's top 100 hit list within each genre. In total, 72 accounts (half male and half female) were assigned to each music genre, and each user account was connected to a programmed script that 'mimicked' the behavior of ordinary Spotify users. This included signing in to the service, playing 10 selected songs, and signing out again, according to the following schedule:

8 am: group 1 (12 bots/genre, 48 bots in total)  
9 am: group 2 (12 bots/genre, 48 bots in total)  
10 am: group 3 (12 bots/genre, 48 bots in total)  
11 am: group 4 (12 bots/genre, 48 bots in total)  
12 pm: group 5 (12 bots/genre, 48 bots in total)  
1 pm: group 6 (12 bots/genre, 48 bots in total)

After each session, a script documented the music recommendations provided to the users in Spotify's Discover section. Because of their programmed nature, the bot users never made any mistakes (such as clicking on the wrong link, or accidentally skipping a song), which is very different from human users. We do not know if the Spotify client could sense their programmed nature, but we received no indications that it did. Using randomized behavioral patterns for the bots could have decreased their "robotic" conduct, but it also would have made comparisons between users much more difficult. Hence, we decided to stick to a

controlled experimental setup. In total, data was collected once per day for each user, during four days between June 18 and 22, 2016. The data was collected using 10 virtual Windows computers and was saved in the shape of screenshots and html data. We conducted the analysis using Microsoft Excel and Google Spreadsheets.

As a whole, our experiment bore some similarity to reverse engineering, a strategy that aims to figure out how technology works by back-tracking its outputs. By studying what kind of music recommendations Spotify delivered to pre-designed users, we were hoping to understand more about patterns in the system's outputs. It has been pointed out, however, that reverse engineering comes with many problems, such as the inability to say something about the cultural work that lies behind the system, and approaching the digital sphere as something stable whose inner workings can be fully discerned (Seaver 2014, see also Introduction and Snickars in this issue). We want to stress that although our study is partly informed by reverse engineering methodologies, our primary interest was not to uncover any presumed 'secrets' in Spotify's music recommendations. Therefore, our analytical approach focuses less on how recommendations function and why, and more on *what they do* in the world.

Knowing that there would always remain blank spots and inconsistencies in the data (since digital technologies are inherently slippery and mutable), we also approached the Spotify service as a "black box" in the classic cybernetic sense of the term (see Pickering 2011: 21).<sup>7</sup> This involves seeing black boxes as inherently ungraspable and ubiquitous, rather than as technical systems that might become fully transparent to our understanding. In this way, our approach to Spotify's music recommendation system may be described as a process of "tinkering", rather than a strict and rule-bound scientific experimentation. In a "tinkering" spirit, the process of arriving at the results in this article were also far from linear and involved many detours, adjustments, and reconsiderations.<sup>8</sup>

### **Tracking Gendered Streams: Results from a Bot Experiment**

The focus of our analysis has been the supposedly personalized music recommendations delivered as "Top Recommendations for You" in Spotify's Discover section.<sup>9</sup> "Top Recommendations for You" is the first content category that users meet when browsing this section, and it can therefore be seen as particularly significant in terms of positioning users and producing meaning around music. The analysis was limited to recommended artists, which means that we did not consider whether users had been recommended different songs or albums by the same artist. Furthermore, since data was collected once per day and Spotify's music recommendations did not update on a daily basis, many artists appeared multiple times for the same user. Such duplicates were removed from the

data set. In sum, our analysis was based on 492 different artist recommendations that were displayed to our bots during the course of the experiment.

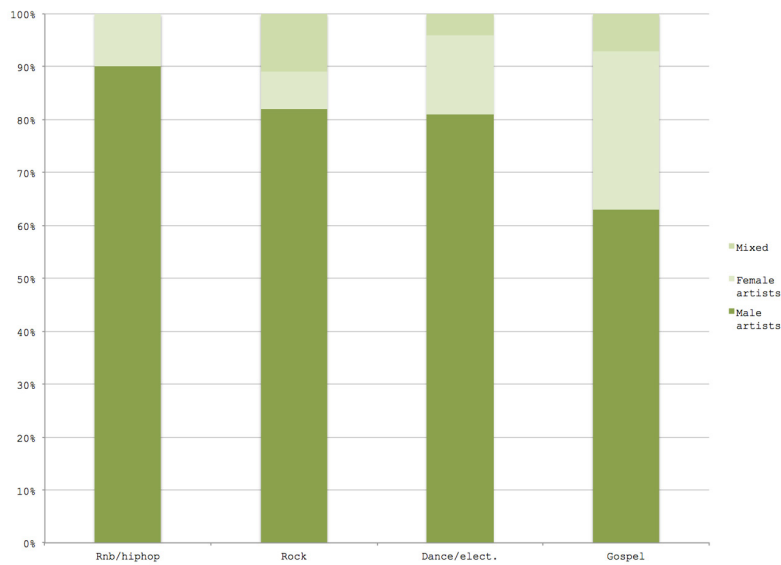
In the first part of the analysis, we explored the extent to which male and female registered users within each music genre had been given the same artist recommendations. This would tell us whether Spotify's recommendation algorithms seemed to assume that our male and female bots had the same taste in music or not. If male and female bots in each genre were given identical recommendations, it could be inferred that the recommendation algorithms had not treated them differently. If the bots were given non-matching music recommendations, however, we would be able to say that within the scope of this particular experiment, Spotify's music recommendation system seemed to respond to the registered gender of the bots.

The results of this analysis showed that 86 percent of the rnb/hiphop bots, 93 percent of the rock bots, 93 percent of the gospel bots, and 78 percent of the dance/electronic bots had largely been recommended music by the same artists, irrespective of their registered gender. The remaining bots had been given very different sets of recommendations. We call such bots "outliers," and the extent to which these bots were male or female differed marginally between the music genres. More specifically, four female and six male rnb/hiphop bots, one female and four male rock bots, three female and two male gospel bots, and nine female and seven male dance/electronic bots were defined as outliers. In other words, there were small indications that our male and female registered bots had been differently targeted as outliers, but it would be precarious to draw any conclusions from this result since the total number of outliers was so small. Interestingly, however, the dance/electronic bots received an overall higher percentage of outlier recommendations than the bots in the other genres.

On the other hand, the analysis showed that a few specific artists were recommended to slightly more male than female registered users, or vice versa. Such seemingly gender skewed recommendations were most common in the rock genre. Here, for instance, 19 of the female bots, but only 12 of the male bots had been recommended the "poppy, jittery, upbeat, math rock/post-punk sound" of the British all-male band Foals.<sup>10</sup> Similarly, the all-male band The Neighborhood, mixing "atmospheric indie rock, electronica, and hip-hop beats with melodic R&B-inflected vocals", was recommended to 32 male bots, as compared to 27 female.<sup>11</sup> We found several examples of slightly gender skewed artists in the rnb/hiphop and dance/electronic genres too,<sup>12</sup> but not in the gospel genre.

In the second part of the analysis we explored the gender ratio among the musicians recommended to our bots. For this purpose, we tagged every artist recommendation according to the gender presentation of the artist. Our reading of gender was based on one or several of the following elements: the pronoun

used in texts about the musician(s), the name of the artist or band members (artist name or personal name), and/or photographs of the artist or band. If a duo or band consisted of both male and female artists, they were tagged as “mixed.”<sup>13</sup> We did not find any musician who explicitly positioned themselves outside the gender binary. By using the criteria above, we were able to define the gender of 485 (or 99 percent) of the unique artist recommendations given to our bots.<sup>14</sup>



**Table 1.** Percentage of recommendations for male artists, female artists, or mixed bands (irrespective of bot gender). The percentages correspond to the following number of artist recommendations: Rnb/hiphop (1730), Rock (2246), Dance/Electronic (1682), Gospel (1644).

Out of these 485 artists and bands, 386 (or 80 percent) were identified as male, and 73 (or 15 percent) were identified as female. 24 (or 5 percent) were tagged as mixed duos or groups. Thus, male artists were highly overrepresented in Spotify’s music recommendations during the course of the experiment. Since Spotify does not publicly announce (or perhaps even register) the self-attributed gender of their artists, we do not know if these figures are representative of the overall gender proportions of available artists on the service. But in any case, our figures revealed that a vast majority of the artists recommended during this particular experiment performed as male. When investigating whether the registered gender of our bots was related to the gender presentation of the artists recommended to them, we found that although male and female users in each genre were given some non-matching artist recommendations, the proportions of male artists, female artists and mixed bands were almost identical.

Table 1 shows the percentage of recommendations for male artists, female artists, and bands with both males and females within each music genre when duplicates between bots were included—figures which, again, were almost identical when the data was broken down into male and female users. Neither did we find any significant differences between outliers and other bots. As the table demonstrates, rnb/hiphop was the genre with the highest dominance of male artist recommendations (90 percent), followed by rock (82 percent), dance/electronic (81 percent), and gospel (63 percent). The rnb/hiphop users were the only group of bots that did not receive any recommendations concerning bands with both male and female artists.

The gospel bots received the highest percentage of recommendations of female artists (30 percent), as well as the second largest percentage of bands with both male and female musicians. The rock and dance/electronic bots were recommended almost the same share of male artists and bands (82 percent and 81 percent respectively). However, the rock bots had a higher percentage of mixed bands in their recommendations (7 percent female artists, and 11 percent mixed duos/groups), while the dance/electronic bots received more female artist recommendations (15 percent), than mixed groups (4 percent).

To summarize, our analysis indicated that overall, Spotify's music recommendation system had not treated our male and female bots differently. In fact, between 78 and 93 percent of the male and female bots in each music genre were given nearly identical recommendations. The remaining percentages were made up of users that had been given a large number of outlier recommendations. While the tendency for users to be positioned either as 'outliers' or as adhering to the genre norm is an interesting finding in itself (and a topic of discussion which we will get back to shortly), we could not find any significant indications that this was related to the registered gender of the users.

Further, our analysis did not indicate that Spotify's music recommendation algorithms assumed our male and female registered bots to have different preferences regarding the gender of artists. Instead, male and female bots within each music genre were largely recommended the same percentage of male artists, female artists and mixed duos or bands. However, the analysis did show that Spotify's music recommendations were heavily geared towards recommending music by male artists to all users during the experiment. If our bots would have continued to listen to the music they were recommended, between 63 and 90 percent of their musical intake would have come from male artists (depending on genre). In the genre with the least female artists (rock), only 7 percent of the recommendations concerned female artists.

## Discussion: Gendering Music Streams

To some extent, the results of our experiment support the notion that popular music is a gendered phenomenon. However, the gendering of Spotify users and the tracking of gendered streams have shown to be neither straightforward nor unambiguous. Jumping back to the initial stages of our experiment, the requirement to register gender—male or female—when signing up for the service was itself a precondition for this study. The mandatoriness of taking up a gendered position, together with the compulsory self-identification within the confines of binary gender, can be seen as an indication that user profiling based on gender was considered vital to the functioning of the software and its recommendation system, or that such profiling was central to the company’s monetization strategies (Bivens 2015). In either case, following Butler (2004), the mandatory gender registration illustrates how identification as either male or female is a performative act, necessary for the production of intelligible subjectivity in user interaction with the service. An obvious effect of the requirement to take up such narrow “menu-driven identities” (Nakamura 2002) is that people identifying outside the gender binary have to either abstain from using the service, or choose to misrepresent themselves—issues which, for several years, have spurred criticism in the Spotify user community.<sup>15</sup>

During the fall of 2016, Spotify opened up for additional forms of self-identification by adding a third option to their mandatory gender field, now consisting of “male”, “female”, and “nonbinary”. This was in line with developments seen in other services, such as Facebook, where the launch of custom gender options in 2014 gained much media attention (Bivens & Haimson 2016). However, as Bivens (2015: 6) demonstrate in her study of the Facebook API, “deep in the database, users who select custom gender options are re-coded—without their knowledge—back into a binary/other classification system that is almost identical to the original 2004 database storage programming”. This, Bivens argues, is a way of simultaneously serving users’ need for genderqueer identifications, and offering advertising clients “a more marketable and ‘authentic’ (yet, paradoxically, misrepresented) data set” (Bivens 2015: 7). Because Spotify’s third gender option was introduced after the finalization of our data collection, we have not been able to interrogate its materialization in code or its effects on music recommendations; these will be important questions for future studies. We note, however, that Spotify first rolled out the nonbinary feature in a few select countries (including Sweden, Australia, UK and the US),<sup>16</sup> with countries such as Brazil, Canada, France, Germany, Italy, Japan and Mexico still being limited to the binary options at the time of writing this article. Furthermore, gender identification (within or outside the binary) is still compulsory, which suggests that gender data is even now considered critical to the service’s profiling of users.

While the compulsory, menu-driven identification brings gendered subjects into existence on the service, it is not immediately clear whether and how gender is continuously enacted in the following interplay between users and the algorithmic system. Our case study demonstrated that, overall, the bots in each music genre had received the same music recommendations, regardless of their self-attributed gender. In selected cases, we detected minor differences in terms of the extent to which our male and female bots were positioned as outliers, and in the extent to which our male and female bots had been targeted with specific artist recommendations. However, these differences were very small, and it cannot be known whether they are actually a consequence of algorithms responding to the initial gender presentation of our bots, randomness, or results of beta-testing a new system.

In contrast to the mandatoriness of gender identification at registration, this absence of gender-specific recommendations illustrates that software solutions may carry the potential to move beyond essentializing notions of identity. As John Cheney-Lippold (2011) notes, there lies a progressive potential in how algorithmic contexts construct identity categories such as ‘male’ or ‘female’ as neither fully self-selected nor “determined by one’s genitalia or even physical appearance” (ibid: 165). Instead, categories are flexible and fluid, continuously inferred upon individuals based on their practices and doings, as compared to other individuals’ practices and doings (cf. Bivens & Haimson 2016). In our case, despite their self-attributed gender, most bots in each genre seemed to be constituted as similar because they all listened to the same music in the same way.

Such constant feedback loops of user behaviors and algorithmic content filterings can be said to accentuate the performative character of the service as well as the ways in which identity categories become open to negotiation (Cheney-Lippold 2017; Kitchin 2017). Because users are requested to engage in continuous acts of music selection and deselection, they can also challenge and transform any normative expectations that might come with these requests (cf. Butler 2004). This may be done through intentional acts of resistance, such as misrepresenting one’s gender when signing up for the service. It may also happen when a user intentionally or unintentionally ignores the prescribed content. In either case, Spotify’s system for delivering music recommendations is not a closed entity which inescapably steers user behavior, but a system that is open to acts of contestation at the front-end, as well as to development and transformations at the back-end. In other words, algorithms “are never fixed in nature, but are emergent and constantly unfolding” (Kitchin 2017: 21).

While the performative and citational nature of recommendation systems may allow us to move beyond essentializing ideas of binary gender, such systems might also—as Cheney-Lippold (2017) has noted—lead to new forms of dynamic,



statistical stereotyping based on behavior rather than demographic categories. This was indicated by the fact that bots listening to different genres also received different recommendations, although our study did not give detailed information about how this dynamic categorization worked. However, every user will in some sense always feed conventionally gendered data into the Spotify system, because they can only make themselves known to the service as conventionally gendered subjects (male, female, or—in select cases—nonbinary).

One additional and unexpected result of our analysis was that some users were constructed as outliers in terms of their music recommendations. While the meaning of such odd user profiles could arguably be interpreted in different ways, we suggest that outlier recommendations position certain users as less mainstream, and more niche and exploratory than others in their music taste. We could find no clear signs that the users' self-attributed gender coincided with outlier status, but we did notice differences between the genres. Dance/electronic was the genre with most outlier bots (22 percent), as compared to rnb/hiphop (14 percent), and rock/gospel, where only 7 percent of the bots had been treated as outliers. This suggests that dance/electronic fans are more often treated as exploratory and adventurous in their music taste, as compared to their rock and gospel counterparts. Given that musical connoisseurship, expertise and agency are characteristics that have frequently been associated with masculinity (eg. Straw 2013; Werner and Johansson 2016: 187), such differences between genres might be said to reinforce the gendering of music styles, possibly constructing dance/electronic—and to some extent rnb/hiphop music—as more niche, exploratory and hence masculine genres than rock and gospel. This is in contrast to some traditional co-constructions of gender and music style, such as the positioning of rock music as masculine (Railton 2001; Bannister 2006; Whiteley 2013).

The most significant result of the study concerned the extent to which our bots had been recommended music by either male or female artists.<sup>17</sup> Our study revealed an overwhelming majority of male artist recommendations in all four music genres. This comes as no surprise, given that the music industry has long been understood as a male dominated domain and, hence, a field in which other gender positions are marginalized (Leonard 2007; Cohen 2013). Still, we find the results remarkable. While there were exceptions to the construction of male-as-norm among musicians, male-defined artists were indisputably privileged during the course of our experiment. It should again be noted that music recommendations provided by algorithms operating under the hood of the Spotify client are created and developed in a larger context (cf. Seaver 2013; Kitchin 2017) and also work together with a range of other cultural intermediaries such as record labels, music aggregators, and music critics. Spotify is thus not an isolated actor in generating gendered streams, but the service did appear to contribute to the construction

of music production as a domain of masculinity. Indeed, the privileging of male artists could be read as producing this streaming service as a masculine context in itself.

Such gendering of music in general, and of the Spotify service in particular, has representational and material effects for both fans and musicians, which points to another aspect of the performative power of the recommendation system. At a symbolic level, the gender representation in artist recommendations encourages specific ways of defining and recognizing musical success. For instance, few female artist recommendations imply fewer opportunities for imagining music talent as a property of femininity—for fans as well as for artists. Notably, the skewed representation might also have material effects in the sense that male artists receive greater financial compensation. Spotify is uniquely positioned to ensure more plays for artists through selective exposure and promotion. But as our study demonstrates; this curatorial authority was deployed in ways that most likely maintained male material privileges in the music industries. Both a consequence and a cause, our bots were urged to financially support and sustain the fame of male musicians. Thus, they were requested to take part in a particular construction of binary gender as well as its power differentials.

## **Concluding Remarks**

The digital methods used in this experiment has enabled an analysis of Spotify's recommendation system during specific sampled circumstances, which brings with it certain limitations. Most notably, the algorithmic structures behind music recommendations continuously change due to developer decisions as well as to the feedback loops that adjust the system's outputs (see e.g. Cheney-Lippold 2011, Seaver 2014). This means that our study is not necessarily replicable or generalizable, as the software system we were engaging with in June 2016 was most likely very different from today. Nevertheless, our study has provided insights into the ways in which the interplay between Spotify's recommendation system and its users is a performative process through which user identities are continuously produced and enacted. Importantly, however, such performative processes do not necessarily involve the construction or reinforcement of gendered music preferences. As demonstrated in this article, Spotify did not appear to infer gendered taste profiles on our bot users. In a majority of the cases, our male and female bots were given identical music recommendations. What the study did show, however, was that Spotify had displayed an equal tendency to recommend male artists to both our male and female bots. In extension, this implies that our users were prompted to take part in the iterative co-production of male-as-norm in the music industries, thereby also reproducing hegemonic gender conventions of masculine artistry and fame.

As a last remark, we want to point out that music recommendations only represent one of the ways in which gender matters on Spotify, and other aspects—such as advertising strategies—may be even more important for how gender materializes on the service. As Bivens (2015) has noted, the mandatoriness of gender registration indicates that gender profiling is essential, either to the functioning of the software or to the service’s monetization strategies. Because gender profiling did not seem to have an immediate impact on music recommendations, one might speculate whether it is instead critical for advertising purposes.

Moreover, gender is not the only user-provided identification that is requested upon registration and thus have the potential to affect music recommendations. Other identity markers such as age, cell-phone type, or country of residence may be worthy of attention and can possibly intersect with gender performances in complex ways. In future research, then, we believe that key insights could be gained by analysing gender alongside other social categories made relevant in algorithmic user profiling. Relatedly, ethnographic studies of software design processes and the rationales behind developers’ design decisions (e.g. Seaver 2014), as well as research on the perceptions and practices of streaming service audiences (e.g. Nylund Hagen 2015; Werner & Johansson 2016), provide important contextualizations for our study. However, while this type of research contributes to a broader understanding of the gendered dimensions of Spotify’s music recommendations, we suggest that scholars in the humanities and social sciences also need to experiment with new ways of engaging with and knowing about digital services. Bot methods might—as we have hopefully shown—provide one such opportunity.

**Maria Eriksson** is a PhD candidate in Media and Communication Studies and an affiliated researcher at Humlab, Umeå University. Her background is in Social Anthropology, with a special interest geared towards digital and economic anthropology. E-mail: [maria.c.eriksson@umu.se](mailto:maria.c.eriksson@umu.se)

**Anna Johansson** is an Associate Senior Lecturer in Ethnology at HUMlab, Umeå University. After receiving her PhD in 2010, her research has focused on social media and gender, online health activism and processes of politicization, and the body in digital culture. Major research interests are digital ethnography and qualitative methods. E-mail: [anna.johansson@umu.se](mailto:anna.johansson@umu.se)

## Notes

i Some of the digital methods used in this article are non-compliant with Spotify's Terms of Service (ToS). The data collection has ended and did not involve any user data. With the public and academic interest in mind, we appreciate Spotify's forbearance with any trespassings of ToS that our data collection involved.

<sup>1</sup>While it is important to point out that heteronormativity (or the entwinement of sex, gender and desire) is constitutional for binary gender in Butler's work, our study focuses mainly on the gender dimension.

<sup>2</sup>For a similar methodological approach, see for example Boshmaf et al (2011), Eriksson and Johansson forthcoming, and Snickars this issue. Also see the work of Christian Sandvig et. al, available at <http://auditingalgorithms.science/> (accessed 10/10/2017).

<sup>3</sup>For a similar approach to experimental research in the humanities and social sciences, see for example Nigel Thrift's (2008: 12) discussions on 'playful' experimental methods.

<sup>4</sup>Arguments which are similar to Lamere's have also re-surfaced on Spotify's own blog, see <https://insights.spotify.com/se/2014/02/10/men-and-women-as-music-fans/> (accessed 27/10/2016).

<sup>5</sup>Limiting the study to 288 bots primarily had to do with hardware restrictions. More users would require more processing power than we had access to.

<sup>6</sup>The gender neutral option is discussed in this post: <https://community.spotify.com/t5/Implemented-Ideas/Make-a-gender-neutral-option-for-profile-sign-up/idi-p/482938> (accessed 11/04/2017). Because our data collection was finalized in June 2016, however, we were not able to include this third gender category in our set-up. This calls for future research and demonstrates the precariousness in studying digital services that are constantly subjected to updates and modifications – an issue which we will return to in the final discussion.

<sup>7</sup>The concept of the "black box" has been used in many ways – not least within the field of Actor Network Theory, and in the works of Bruno Latour. Here, however, we solely wish to denote the cybernetic use of the term as laid out by Pickering (2011).

<sup>8</sup>Before the data presented in this article was collected, we had done two (less successful) pre-studies. Initially, we had troubles establishing a stable system of data collection. During the second pre-study for instance, a majority of the bots were not given any music recommendations at all, and we were unable to find out exactly why. This shows how difficult it is to engage with algorithmic systems whose operational logics are hidden.

<sup>9</sup>The remaining categories of music recommendations that were not included in the analysis include "Discover Weekly," "New Releases For You," and "Because You Listened to XXX..."

<sup>10</sup>See <https://open.spotify.com/artist/6FQqZYVfTNQ1pCqfkwVFEa> (accessed 09/05/2017)

<sup>11</sup>See <https://open.spotify.com/artist/77SW9BnxLY8rJ0RciFqkHh/about> (accessed 10/10/2017).

<sup>12</sup>For example, in the rnb/hiphop group, none of the female bots, but six of the male rnb/hiphop bots had been recommended music by the artist J.R., and in the dance/electronic group, six female and only one male bot had been recommended music by Robin Thicke, one of the most “charismatic, flashy and commercially successful R&B acts of the 2000s and 2010s”, according to Spotify. See <https://open.spotify.com/artist/0ZrpamOxcZybMHGg1AYtHP> (accessed 14/05/2017)

<sup>13</sup>While we recognize that this is a problematic undertaking and that there might be cases where our assumptions might not match the self-identification of artists, we believe that pronouns, names and images are fairly established ways of performing and reading gender – and thus relevant criteria for our purposes.

<sup>14</sup>The remaining 1 percent has been excluded from the analysis. These were artists that we could not find any information about online.

<sup>15</sup>See e.g. <http://mxactivist.tumblr.com/post/69298310891>, <https://community.spotify.com/t5/Live-Ideas/Make-a-gender-neutral-option-for-profile-sign-up/idi-p/482938>, <https://community.spotify.com/t5/Help-Accounts-and-Subscriptions/Are-non-binary-people-not-welcome-to-use-Spotify/m-p/1173171/highlight/true#M160172>, <http://mxactivist.tumblr.com/post/69768251170> (all accessed 13/04/2017)

<sup>16</sup>See <http://mxactivist.tumblr.com/post/150489712620> (accessed 11/04/2017).

<sup>17</sup>An alternative interpretation could for example be that users who receive outlier recommendations are in much greater need for special musical guidance, and hence are perceived as less independent. Looking at the outlier recommendations, however, many seemed to be artists who are lesser known. This, we argue, indicates that the outlier recommendations are geared towards a more specialized and exploratory taste in music.

## References

- Amoore, Louise & Piotukh Volha (eds.) (2016): *Algorithmic Life: Calculative Devices in the Age of Big Data*, London and New York: Routledge.
- Baker, Paul & Amanda Potts (2013): “‘Why do white people have thin lips?’ Google and the perpetuation of stereotypes via auto-complete search forms”, *Critical Discourse Studies*, 10:2, 187–204.
- Bannister, Matthew (2006): *White boys, white noise: masculinities and 1980s indie guitar rock*, London: Ashgate.
- Bayton, Mavis (2013): “Women and the electric guitar,” Sheila Whiteley (ed.): *Sexing the Groove: Popular Music and Gender*, New York: Routledge.
- Beer, David (2013): *Popular Culture and New Media: The Politics of Circulation*, London and New York: Palgrave Macmillan.
- Berggren, Kalle (2013): “Degrees of Intersectionality: Male Rap Artists in Sweden Negotiating Class, Race and Gender,” *Culture Unbound*, 5, 189–211.
- Berry, David M. (2011): *The Philosophy of Software: Code and Mediation in the Digital Age*, London and New York: Palgrave Macmillan.
- Bivens, Rena (2015): “The gender binary will not be deprogrammed: Ten years of coding gender on Facebook,” *New Media & Society*. Published online ahead of

- print: <http://journals.sagepub.com/doi/abs/10.1177/1461444815621527> (accessed 10/04/2017).
- Bivens, Rena & Oliver L Haimson (2016): "Baking Gender into Social Media Design: How Platforms Shape Categories for Users and Advertisers," *Social Media + Society*, 2:4, 1–12.
- Boshmaf, Yazan, Ildar Muslukhov, Konstantin Beznosov, & Matei Ripeanu (2011): "The Socialbot Network: When Bots Socialize for Fame and Money," *ACSAC*, Dec 5–9: <http://dl.acm.org/citation.cfm?id=2076746> (accessed 10/04/2017).
- Butler, Judith (1990): *Gender trouble: feminism and the subversion of identity*, New York: Routledge.
- Butler, Judith (2004): *Undoing Gender*. New York: Routledge.
- Butler, Judith (2010): "Performative Agency," *Journal of Cultural Economy*, 3:2, 147–161.
- Cheney-Lippold, John (2011): "A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control," *Theory, Culture & Society*, 28:6, 164–181.
- Cheney-Lippold, John (2017): *We are Data: Algorithms and the Making of our Digital Selves*, Ne York: NYU Press.
- Choi, Grace Y (2016): "'Who Run the Music? Girls!' Examining the Construction of Female Digital Musicians' Online Presence," *Popular Music and Society*, 1–14, <http://dx.doi.org/10.1080/03007766.2016.1174419>.
- Cohen, Sara (2013): "Men making a scene: rock music and the production of gender," Whiteley, Sheila (ed.): *Sexing the Groove: Popular Music and Gender*. New York: Routledge.
- Dixon-Roman, Ezekiel (2016): "Algo-Ritmo: More-Than-Human Performative Acts and the Racializing Assemblages of Algorithmic Architectures," *Cultural Studies Critical Methodologies*, 16:5, 482–490.
- Drew, Rob (2005): "Mixed Blessings: The Commercial Mix and the Future of Music Aggregation," *Popular Music and Society*, 28:4, 533–51.
- Eriksson, Maria & Anna Johansson (forthcoming): "'Keep Smiling!': Time, Functionality and Intimacy in Spotify's Featured Playlists," forthcoming in *Cultural Analysis, Special Issue on The Inheritance of the Digital: Ethnographic Approaches to Everyday Realities In, Of, And Through Digital Technologies*.
- Fatherstone, Mike (2006): "Archive," *Theory, Culture & Society*, 23:2–3, 591–96.
- Frith, Simon & Angela McRobbie (1991): "Rock and sexuality," Frith, Simon & Andrew Goodwin (eds.): *On Record: Rock, pop and the written word*. London/New York: Routledge, 371–389.
- Fuller, Matthew (2003): *Behind the blip. Essays on the culture of software*, Brooklyn, NY: Autonomedia.
- Gadir, Tami (2016): "Resistance or Reiteration? Rethinking Gender in DJ Cultures," *Contemporary Music Review*, 35:1, 115–129.
- Galuszka, Patryk (2015): "Music Aggregators and Intermediation of the Digital Music Market," *International Journal of Communication*, 9, 254–73.
- Gavanas, Anna & Rosa Reitsamer (2013): "DJ technologies, social networks and gendered trajectories in European DJ cultures", Attias, Bernardo, Anna Gavanas & Hillegonda C Rietweld (eds.): *DJ culture in the mix: Power, technology, and social change in electronic dance music*, New York: Bloomsbury Academic, 51–77.
- Gillespie Tarleton (2014) "The Relevance of Algorithms", Tarleton Gillespie, Pablo J. Boczkowski & Kirsten A. Foot (eds.) *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge & London: MIT Press, 167–194.
- Haraway, Donna (1997): *Modest Witness@Second Millennium.FemaleMan@Meets\_OncoMouse: Feminism and Technoscience*. New York: Routledge.
- Hayles, Katherine (1999): *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics*. Chicago: University of Chicago Press.
- Introna, Lucas D & Helen Nissenbaum (2000): "Shaping the Web: Why the Politics of

- Search Engines Matters,” *The Information Society*, 16:3, 169-185.
- Jeffries, Michael P (2011): *Thug life: race, gender, and the meaning of hip-hop*, Chicago: Chicago University Press.
- Kitchin, Rob (2017): “Thinking Critically about and Researching Algorithms,” *Information, Communication & Society*, 20:1, 14-29.
- Kjus, Yngvar (2016): “Musical exploration via streaming services: The Norwegian experience,” *Popular Communication*, 14:3, 127–136.
- Lamere, Paul (2014): “Gender Specific Listening,” *Music Machinery*, Feb 10: <https://musicmachinery.com/2014/02/10/gender-specific-listening/> (accessed 10/04/2017)
- Leonard, Marion (2007): *Gender in the Music Industry: Rock, Discourse and Girl Power*, Aldershot, England: Ashgate.
- Maasø, Arnt (2017): “Music Streaming, Festivals, and the Eventization of Music,” *Popular Music*, 1-22, <https://doi.org/10.1080/03007766.2016.1231001>.
- Mackenzie, Adrian (2006): *Cutting Code: Software and Sociality*, New York: Peter Lang.
- Mager, Astrid (2012): “Algorithmic Ideology,” *Information, Communication & Society*, 15:5, 769–787.
- Manovich, Lev (2013): *Software Takes Command*, New York: Bloomsbury.
- Modell, Amanda (2015): “Mapping the Music Genome: Imaginative Geography in Pandora Internet Radio and the Genographic Project,” *Media Fields Journal*, 10, 2-13.
- Morris, Jeremy W (2012): “Making Music Behave: Metadata and the Digital Music Commodity,” *New Media & Society*, 14:5, 850–66.
- Morris, Jeremy W (2015): *Selling digital music, formatting culture*, Oakland, California: University of California Press.
- Nakamura, Lisa (2002): *Cybertypes: race, ethnicity, and identity on the Internet*, New York: Routledge
- Nylund Hagen, Anja (2015): “The Playlist Experience: Personal Playlists in Music Streaming Services,” *Popular Music and Society*, 38:5, 625–645.
- Olofsson, Jennie (2015): ““Did you mean: why are women cranky?”: Google – a means of inscription, a means of de-description?,” in Svensson Patrick & David Theo Goldberg (eds.), *Between humanities and the Digital*, Cambridge, MA: MIT Press.
- Parisi, Luciana (2013): *Contagious architecture: computation, aesthetics, and space*, Cambridge, MA: MIT Press.
- Pickering, Andrew (2011): *The Cybernetic Brain: Sketches of Another Future*, Chicago & London: University of Chicago Press.
- Pough, Gwendolyn D (2004): *Check It While I Wreck It: Black Womanhood, Hip-Hop Culture, and the Public Sphere*, Boston: Northeastern University Press.
- Railton, Diane (2001): “The gendered carnival of pop,” *Popular Music*, 20:3, 321–331.
- Rogers, Richard (2013): *Digital methods*, Cambridge, MA: The MIT Press.
- Ruppert, Evelyn, John Law & Mike Savage (2013): “Reassembling Social Science Methods: The Challenge of Digital Devices,” *Theory, Culture & Society*, 30:4, 22–46.
- Seaver, Nick (2014): “On Reverse Engineering: Looking for the Cultural Work of Engineers,” *The Atlantic*, Jan 28: <https://medium.com/anthropology-and-algorithms/on-reverse-engineering-d9f5bae8712> (accessed 10/04/2017).
- Seaver, Nick (2013): “Knowing Algorithms,” *Media in Transition* 8, Cambridge, MA: <http://nickseaver.net/papers/seaverMIT8.pdf> (accessed 13/04/2017).
- Straw, Will (2013): “Sizing up record collections: gender and connoisseurship in rock music culture,” Sheila Whiteley (ed.): *Sexing the Groove: Popular Music and Gender*, New York: Routledge.
- Sundén, Jenny (2015): “On trans-, glitch, and gender as machinery of failure,” *First Monday*, 20:4: <https://doi.org/10.5210/fm.v20i4.5895> (accessed 13/04/2017)
- Thrift, Nigel (2008): *Non-Representational Theory: Space, Politics, Affect*, London & New York: Routledge.

- Wajcman, Judy (2007): "From Women and Technology to Gendered Technoscience," *Information, Communication & Society*, 10:3, 287–298.
- Werner, Ann & Sofia Johansson (2016): "Experts, dads and technology: Gendered talk about online music," *International Journal of Cultural Studies*, 19:2, 177–192.
- Whiteley, Sheila (ed.) (2013): *Sexing the Groove: Popular Music and Gender*, New York: Routledge.
- Whiteley, Sheila & Jennifer Rycenga (eds.) (2006): *Queering the popular pitch*, New York: Routledge.





## More of the Same – On Spotify Radio

By Pelle Snickars

### Abstract

Spotify Radio allows users to find new music within Spotify's vast back-catalogue, offering a potential infinite avenue of discovery. Nevertheless, the radio service has also been disliked and accused of playing the same artists over and over. We decided to set up an experiment with the purpose to explore the possible limitations found within "infinite archives" of music streaming services. Our hypothesis was that Spotify Radio appears to consist of an infinite series of songs. It claims to be personalised and never-ending, yet music seems to be delivered in limited loop patterns. What would such loop patterns look like? Are Spotify Radio's music loops finite or infinite? How many tracks (or steps) does a normal loop consist of? To answer these research questions, at Umeå University's digital humanities hub, Humlab, we set up an intervention using 160 bot listeners. Our bots were all Spotify Free users. They literally had no track record and were programmed to listen to different Swedish music from the 1970s. All bots were to document all subsequent tracks played in the radio loop and (inter)act within the Spotify Web client as an obedient bot listener, a liker, a disliker, and a skipper. The article describes different research strategies when dealing with proprietary data. Foremost, however, it empirically recounts the radio looping interventions set up at Humlab. Essentially, the article suggests a set of methodologies for performing humanist inquiry on big data and black-boxed media services that increasingly provide key delivery mechanisms for cultural materials. Spotify serves as a case in point, yet principally any other platform or service could be studied in similar ways. Using bots as research informants can be deployed within a range of different digital scholarship, so this article appeals not only to media or software studies scholars, but also to digitally inclined cultural studies such as the digital humanities.

**Keywords:** Spotify Radio, digital methods, music looping, bot intervention, reverse engineering

Snickars, Pelle: "More of the Same—On Spotify Radio", *Culture Unbound*, Volume 9, issue 2, 2017: 184–211. Published by Linköping University Electronic Press: <http://www.cultureunbound.ep.liu.se>

## Introduction – “Same Artists Over and Over”

Sometime during Winter 2014 someone posted a question on the Q&A site Quora: “Why does my Spotify radio sound so repetitive? I feel I am getting a few artists repeated”. At the time, Spotify Radio had been around for more than two years, but Quora users were disappointed. “The radio functionality in Spotify is very crude”, the Finnish ‘info junkie’ Heikki Hietala replied to some frustrated listeners. Maybe Spotify will “come up with something soon, but as for now Spotify Radio is very annoying”. Apparently, Hietala had the same experience of repeated songs being played, and instead recommended the music streaming service Pandora. According to Hietala, the latter had way more successfully “chopped the music up into tiny pieces of metadata [delivering] a truly mesmerising radio function due to the vast amount of information they have on the music” (Hietala 2014).

Quora is a question-and-answer site where users post questions, which are subsequently answered, edited, and organized by the community of users on the same site. Queries on Quora often address tech, which is not surprising since the company was co-founded by two former Facebook employees and is based in Mountain View, California (Google headquarters). Quora also seems to be a site frequented by tech employees themselves, which makes it particularly interesting from a research perspective. Former Tech Lead at Spotify, Erik Bernhardsson, has published almost 30 posts, some with references to discussions on Spotify Radio.

A couple of months after Hietala’s post, another objection in the same vein re-appeared on Quora. In fact, almost identical questions around the inferior functionality of Spotify Radio kept being posted: “How do I get Spotify to stop playing the same few songs for every artist?”; “How do I teach a Spotify radio station to play a wider array of songs?”; “Is the Spotify streaming radio . . . purposefully terrible with the intention of trying to get people to upgrade?” (Quora 2016). Within our ongoing research project on Spotify, we have discussed similar issues around the poor performance of the Spotify Radio algorithm. Naturally, such assumptions reveal a normative claim that the radio algorithm should produce apt recommendations. To answer at least some of these issues, we decided to set up an experiment that would explore Spotify Radio. Essentially, we wanted to uncover why we (usually) didn’t like the songs the radio algorithms suggested we should like. But given normative assumptions about the ways in which Spotify Radio ought to work, the research question was also broader, hinting at the ways in which algorithmic music discovery today features and promotes some artists and simply ignores others. A software driven cultural analyses of music delivery mechanisms could potentially reveal the algorithmic flaws that regulate music recommendations in disfavour of, say, more diverse listening, making less room for emerging musicians or neglected genres (with economic ramifications). Ultimately, an investigation of Spotify Radio would also stress the possible limitations and

restraints found within “infinite archives” of music streaming services.

Our hypothesis was that many streaming services’ radio functions appear to consist of an infinite series of songs. For commercial reasons, Spotify Radio claims to be both personalised and never-ending, yet music seems to be delivered in limited loop patterns. If our hypotheses held true, what would such loop patterns look like? Are Spotify Radio’s music loops finite or endless (given that its algorithm(s) can choose between 30 million songs)? How many tracks (or steps) does a typical loop consist of? Importantly, how is the size of a music loop on Spotify Radio affected by user interaction in the form of likes, dislikes, and skips? Does, for example,  $n$  amounts of likes expand the music loop in terms of novel songs and artists?

This article describes different research strategies and digital methods when dealing with proprietary data as well as the background and the establishment of the radio functionality at streaming services like Spotify. It briefly recounts, for example, what is known about Spotify Radio’s music discovery engine. Foremost, however, the article empirically recounts, discusses, and analyses the radio looping interventions set up at the digital humanities hub, Humlab (Umeå University, Sweden). Essentially, the article suggests a set of methodologies for performing humanist inquiry on (mid-size) big data and black-boxed media services that today increasingly serve as key delivery mechanisms for cultural materials. Spotify serves as a case in point, yet principally any other platform or service could be studied in similar ways.

## Proprietary Data & Research Strategies

One of the users on Quora concerned with the second-rate quality of Spotify Radio was Web designer Bas Leijder Havenstroom: “Why does my Spotify Radio play the same artists over and over for me?” In a re-entry in the thread that followed, he specified what he was puzzled about:

I re-asked this one because this frustrates me . . . Even if I start a radio station based on a playlist with many, many artists, I find that some (specific) artists keep coming back. I have the feeling that this all has to do with commercial reasons. I believe record labels pay Spotify to have their artists to show up in radio stations and random functions more often. (Leyder, Havenstroom 2015)

The major problem in doing contemporary research on streaming music is that claims like the one made by Leijder Havenstroom cannot be tested. The specifics of Spotify’s algorithms are proprietary, and statistics around listener data are a

corporate secret. Some rudimentary data are available via the various Spotify interfaces, ranging from the number of streams for popular songs to artist followers and listeners per month. At the site [yearinmusic.spotify.com](http://yearinmusic.spotify.com) data with the most popular tracks, albums, or playlists (divided, for example, geographically or by time of year) are also available. If a user logs in, personal statistics are displayed in a similar manner. In addition, Spotify occasionally releases listener data and statistics, usually in various tie-ins with magazines or newspapers to gain public attention. Basically, the same strategy (regarding the lack of access to user data) is deployed by other music streaming services. In short, not much data is available.

Consequently, academically there are gaps and lack of knowledge about the ways in which algorithmic music discovery takes place (Kjus 2016). Because of the “lack of transparency in how recommendations and ‘discoveries’ are presented”, as Jeremy Wade Morris and Devon Powers have stated, it is not clear, for example, how promotional messages for artist are featured. At Spotify, an advanced promotional feedback loop today mixes user activity with interface design, and the line between the service as “a distribution outlet” and as a “promotional intermediary” becomes completely blurred (Morris & Powers 2015). Hinting at the metaphorical associations between the streaming service Pandora and Pandora’s box in Greek mythology, Paul Allen Anderson has even claimed that the former connotes “a black box of friendly mysteries” (2015). From a media research perspective, it is simply hard to tell how “music recommendation works – and doesn’t work”, to quote a blog post from Brian Whitman, CTO of the Echo Nest and Principal Scientist at Spotify. Admittedly, computational knowledge and access to data does not automatically lead to profound insights regarding music discovery or the broader relation between culture and tech. Still, people such as Whitman (or Bernhardsson) on the inside do know more (or used to know), but for media researchers it is hard to tell (Whitman 2012). Gaining access to the inside becomes important, or as a former Spotify intern and researcher, Sander Dieleman, stated: “At Spotify, I have access to a larger dataset of songs, and a bunch of different latent factor representations obtained from various collaborative filtering models” (2014). In addition, given that personalisation algorithms alter user experience through interactions with the system, Nick Seaver notes that “it is very difficult, if not impossible for a lone researcher to abandon the subject position of ‘user’ and get an unfiltered perspective” (2014b).

As a result, research undertaken around algorithmic music discovery has been done from a strict technical perspective within the field of computer science (Shao et al. 2009). Computationally-oriented studies have been made regarding, for example, how music recommendations based on artist novelty and similarity work (Lin et al. 2014) or implementing recommendation system based on user behaviour. These studies tend to focus on either applying self-developed algorithmic

mic systems or evolve around mathematical models with, for example, the use of “Gaussian distributions to evaluate each possible genre for the next track” (Yajie & Ogihara 2011). Due to the difficulty (or even impossibility) in obtaining valid data from streaming services, media studies’ perspectives on algorithmic discovery, on the other hand, have tended to favour hermeneutic and critical explorations – i.e., traditional humanistic readings of algorithmic music recommendations (Allen Anderson 2015; Modell 2015; Morris & Powers 2015) or interview-based examinations, sometimes done from an ethnographic or anthropological perspective. Nick Seaver, for example, is currently finishing a long-term anthropological study of developers of music recommender systems (2016). As an anthropologist, he is sceptical about working with digital methods: “While reverse engineering might be a useful strategy for figuring out how an existing technology works, it is less useful for telling us how it came to work that way” (2016). The risk of using digital methods (as reverse engineering), according to him, is that technical details “hidden behind the curtain” become the sole purpose of research. Unlocking corporate secrets, such as how algorithmic discovery works, is not only about technology: “Not everything worth knowing has been actively hidden” (Seaver 2014a).

Seaver’s argument is worth considering. At the same time, he advocates a methodological blind alley, sticking to a traditional and long-established interview-based or participant observation methodology (in his case with software developers). Naturally, such methods can offer valuable insights into the thinking that goes into building algorithms. Then again, within our research project we are mainly interested in using new digital tools to understand the politics of algorithms from a completely different angle. Unlike Seaver, our research project engages in reverse engineering Spotify’s algorithms, aggregation procedures, metadata, and valuation strategies, breaking into the hidden infrastructures of digital music distribution, to study its underlying norms and structures. One point of departure is that Spotify resembles a black boxed service, metaphorically as well as practically (at least from an academic media studies perspective). Another point of departure is that Spotify does not, to put it bluntly, share any data. Lack of access to data today confronts media scholars, (digital) humanists, and social science researchers working with social media studies. “Twitter determines what data are available and how data can be accessed through [their] API”, David Gunnarsson Lorentzen, for example, states in his thesis, *Following Tweets Around* (2016). How Twitter and other similar platforms and services (like Spotify) provide access to data influences how researchers conduct their work: “The central problem [is] that researchers do not know what relevant data are not collected” (Gunnarsson Lorentzen 2016: ii).

Then again, even if Twitter data are biased, research has flourished and benefitted enormously from the ease of access to the relatively open data the servi-

ce provides, at least in comparison to strictly confined music streaming services. Since Spotify user data are not available, the data must be acquired and compiled through other means to perform research, for example, by deploying bots as research informants. As stated, in our experiments we wanted to explore and investigate how Spotify Radio generally functions. What we set out to do in our interventions was to use hundreds of bots to compile user data and essentially reverse engineer what a radio loop at Spotify looks like. As a research strategy, reverse engineering starts with the final and implemented product, in our case Spotify Radio within the streaming service desktop client, and tries to take it apart “seeking clues as to why it was put together in the way it was and how it fits into an overall architecture” (Gehl 2014:10). As an attempt to reveal the procedures of culture and technology at work, reverse engineering can be linked to various forms of hacking practices. Within media studies, reverse engineering has been used both by academic scholars (Friesinger & Herwig 2014) as well as by tech journalist wanting to understand and analyse, for example, how Netflix’s sorting algorithms, vocabulary, and grammar work (Madrigal 2014). In cases where the code remains black-boxed, Rob Kitchin has stated that “a researcher interested in the algorithm at the heart of its workings is left with the option of trying to reverse engineer the compiled software”. Referring to Seaver’s critique, Kitchin has endorsed the use of bots as a research strategy when dealing with proprietary code and inaccessible data. One solution for enhancing clarity is “to employ bots, which posing as users, can more systematically engage with a system, running dummy data and interactions” (Kitchin 2016). This is exactly what we have been doing in our radio looping interventions.

### **A Brief Spotify Radio History**

The Spotify Radio slogan states that the service “lets you sit back and listen to music you love. The more you personalize the stations to match your tastes the better they get” (Spotify Radio 2016). Users of Spotify can start a radio station based on artist, song, playlist, album, or even genre. For users, Spotify Radio is a “lean back experience”, yet with the ability to tune recommendations with thumbs up (like), thumbs down (dislike), or by skipping a song. It is one of many discovery mechanisms or packages associated with music recommendation systems at Spotify, including Discover, Related Artist, Genres, and Moods, Discover Weekly, and lately, Release Radar. In short, the radio functionality allows people (via various algorithms) to find new music within the vast back-catalogue of Spotify, offering a potential infinite avenue of discovery. It is important to stress, however, that the concept of radio has served as the point of departure for the recommendation business within streaming music services. Hence it is essential to study how dif-

ferent radio functionalities have developed, since the radio metaphor for recommendations seems to be diminishing, gradually being replaced by other and more sophisticated modes of machinic suggestions.

According to Noa Resare, Spotify's Free Software Advocate, when a user starts the radio functionality, the Spotify Web client makes a request to "a specialized radio service which holds radio selections for a large subset of all our material" and "[t]he radio suggestions are built using analysis of previous playback data" (2013). Despite the significant amount of data that goes into radio recommendations, for many the feature has been a disappointment. The impression is confirmed if one reads threads and conversations on the Spotify community Web, where the radio functionality is repeatedly criticised. The function, in fact, aroused disappointment right from the start: "Better radio algorithms . . . there are too many repetitions", one user stated already after the launch in 2012 (Lehwark). The critique has remained, and a search using Google autocomplete even suggests the amusing "Spotify radio always [plays the] same song" (if one starts a query for "Spotify radio algorithm").

By and large, the general concept of a personal digital radio station began around 2000. At the time, the idea was introduced to create a separate and individualised radio station for each user depending on personal preferences. In the U.S., within the Music Genome Project, a mathematical understanding of music was developed, which used some 450 attributes to describe and dissect music. Experts tagged songs with different characteristics (genre, instruments, tonality, etc.). Using this information, algorithms organised and bundled music in specified ways. Commercially, the Music Genome Project formed the backend and core technology of Pandora Radio, which after 2000 became the first music streaming and automated recommendation service. Its popularity grew, reaching an estimated 70 million monthly users in 2013. Due to music rights restrictions, however, Pandora Radio is only available in the U.S., Australia, and New Zealand.

For several years, Pandora Radio was widely recognised as the best music recommendation service, and Spotify (all likely) tried to replicate its personalised music offerings. Nonetheless, Spotify seemed to have lacked technological expertise, and as a result started co-operating with the music intelligence company The Echo Nest. In a blog post in December 2011, it was announced that the latter company would now power the new Spotify Radio:

Spotify has over ten million users . . . But there's one thing Spotify didn't have until now: artist and song radio stations. Thanks to Spotify's deal with The Echo Nest, users of the popular music service can now create streaming radio stations based on any artist or song on Spotify. The stations are generated by The Echo Nest's Playlist API, and are

available both to free and premium Spotify subscribers. (The Echo Nest 2011)

This co-operation developed, and in April 2012 Spotify began updating its desktop software with several new features, including a Pandora-like radio station: “Spotify to Take On Pandora With Radio service” (Hachman 2012). Arguably, the Spotify Free model, with ad-supported unlimited streaming access, was especially important in gradually establishing the service in the U.S. An online radio offering, it was proclaimed, “would advance Spotify’s strategy of attracting users with free, ad-supported services who can be converted later into paying subscribers” (Fixmer 2012). At the time, The Echo Nest did not exclusively power Spotify’s radio recommendations. Since its API was open, competitors like Rdio and Deezer were also using it. In March 2014, however, Spotify acquired The Echo Nest, a deal that was said to strengthen its music discovery expertise: “The acquisition supports Spotify’s strategy to grow global music consumption and overall revenue back to the music industry by building the best user experience and music discovery engine for millions of global fans” (The Echo Nest 2014).

According to some press accounts, during the last years Spotify has put The Echo Nest employees in charge of its most important discovery products (Popper 2015). Nevertheless, the music discovery engine at Spotify, technologically sustained by The Echo Nest, remains obscure. Basically, the same uncertainty and unpredictability makes it difficult to research the different algorithms regulating music recommendations on Spotify Radio. In addition, recommendation algorithms vary and the music discovery engine has naturally, like most computational systems, been altered, improved, and updated since its initial release (Chandra 2013). Apparently, the algorithms running the music discovery engine are identical, independent of whether one uses the Free or the Premium service. The only difference is that advertisements play in the Free version which also cannot stream higher audio qualities.

Interestingly, most comments online regarding the ways that music discovery and recommendations at Spotify works refer to a presentation that former Tech Lead, Erik Bernhardsson, gave almost exactly at the time The Echo Nest was purchased. In his talk, Bernhardsson addressed the ways in which Spotify discovery engine functions: “How do we structure music understanding? How do you teach music to machines?” Essentially, Bernhardsson listed five ways to achieve apt recommendations: “Editorial tagging, Audio analysis, Metadata, Natural language processing [and] Collaborative filtering” (Bernhardsson 2014). The last idea determines listeners’ preferences from historical use data, since constant feedback to Spotify is implicit in all streaming behaviour. When someone on Quora asked what data points Spotify Radio uses, Bernhardsson answered that they mainly de-



ployed collaborative filtering through “large scale data mining of user logs . . . to create a statistical model of what artists/albums/track are similar . . . [which is] then post-processed and exposed in the radio service” (Bernhardsson 2013).

It is hard to tell what role collaborative filtering has in relation to Spotify Radio. Most likely different forms of user data analyses are combined. Then again, collaborative filtering is interesting since it is content agnostic. The computational strategy is to look only at user consumption patterns, so the same type of collaborative filtering models can be used to recommend books, films, or music. Since reliance is put on usage data only, popular content will be easier to recommend (compared to ignored content) simply because there is more user data. Collaborative filtering is one way to tap the collective intelligence of Spotify’s millions of users, turning their preferred music and taste into a data layer to personalise everyone’s experience. Then again, as Sander Dieleman has noted, collaborative filtering algorithms specific to music need to pay attention to heterogeneity of content “with similar usage patterns”. As mentioned, it is a difficult method to deploy. Users may, for example, listen to entire albums in “one go, but albums may contain intro tracks, outro tracks, interludes, cover songs and remixes. These items are atypical for the artist in question, so they aren’t good recommendations” (Dieleman 2014, Dieleman et al. 2013).

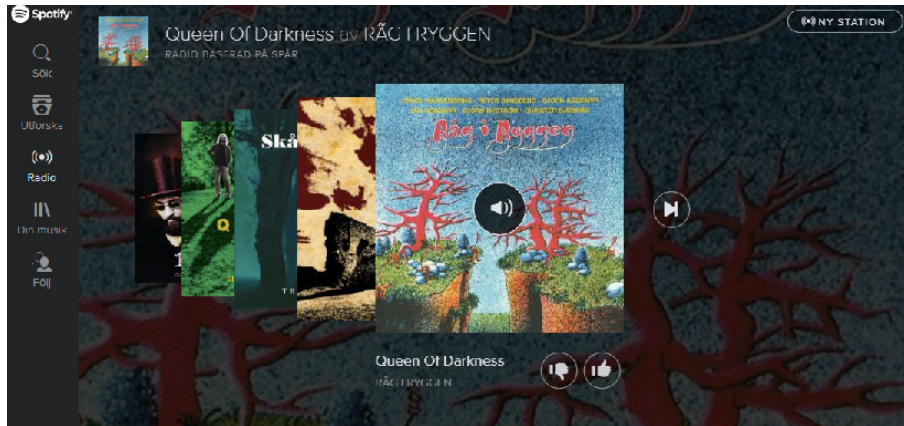
## **Reverse Engineering Radio Loops – Intervention Set Up**

To answer the research questions postulated in the introduction to this article, we and system developers Roger Mähler and Johan von Boer set up an experiment with the purpose to examine several Spotify Radio loops. The loops were constructed using 160 so-called bot listeners. A bot is a small software application that runs automated tasks (or scripts), and we implemented our bots in the Python programming language. Bots appear to be human (at least to the Spotify Web client), which is why they are interesting to use as research informants (Snickars & Mähler 2017).

Our bots were Spotify Free users with literally no track record. They had “heard” no music before they were put into action. We were thus not primarily interested in the personalised recommendations Spotify’s algorithms offered, but rather how Spotify Radio functioned generically. The reason was also practical, as providing our bots with a personal track record would have been hard (if not impossible) to accomplish. In addition, as virtual informants, our bots did not explicitly collect information. They were programmed and designed to search for a track, retrieve subsequent songs, partially interact, and importantly log all data caused by different actions.

As one of our aims with the intervention was to study repetitiveness in loop

patterns, another hypothesis was that the size and structure of radio loops might depend on music genre as well as popularity. Hence, we decided to let our bots “listen” to both a hit song and a less popular track (albeit with some contextual similarity) as our bots would start a radio channel based on Swedish music from the



**Illustration 1.** The Web interface of Spotify Radio with subsequent tracks based on progressive rock band, Råg i Ryggen’s “Queen of Darkness”, including the three characteristic ways of adjusting recommendations: thumbs up, thumbs down, and skip.

1970s. The two 24-hour interventions at Humlab took place in July and September 2016. The bot setup involved a few similar steps. First, bots were named (Anna.01, Cobolt.01, Fred.04.Mercury.04, Jane.12.Lead.01, etc.). Second, the bots –in the form of virtual users acting as research informants within the Spotify Web client – were programmed to “listen” to all subsequent songs that the Spotify Radio algorithm(s) generated. That is, we executed a major and a minor intervention. In the first round, the bots (120 in all, although many failed because of different technological problems) started Spotify Radio based on the highly popular Abba song “Dancing Queen” (released in 1976, with some 65 million streams at Spotify). In the second round, the bots (40 in all, with three failures) used the significantly less popular Swedish progressive rock band Råg i Ryggen’s “Queen of Darkness” (released in 1975, with approximately 10,000 streams) to start a radio channel.

All bots documented all subsequent tracks played in the radio loop, and importantly interacted differently within the Spotify Web client as an “obedient” bot listener, a “liker”, a “dis-liker”, and a “skipper”. These interactions were documented, including tracks and artist played as well as breaks for advertisement. The user scenario (given to the programmers) for the first round of bots generically read as follows:

**User scenario 1:** (approximately 30 obedient listener bots): Starts a radio station based on Abba's "Dancing Queen". Bots passively listen to the full loop. Run time 24 hours. If the radio loop stops playing, bots should be prepared to restart.

**User scenario 2:** (approximately 30 liker listener bots): Starts a radio station based on Abba's "Dancing Queen". Bots like every fifth song. Run time 24 hours. If the radio loop stops playing, bots should be prepared to restart.

**User scenario 3:** (approximately 30 disliker listener bots): Starts a radio station based on Abba's "Dancing Queen". Bots dislikes every fifth song. Run time 24 hours. If the radio loop stops playing, the bots should be prepared to restart.

**User scenario 4:** (approximately 30 skipper listener bots): Starts a radio station based on Abba's "Dancing Queen". Bots skip every fifth song. Run time 24 hours. If the radio loop stops playing, the bots should be prepared to restart.

To introduce randomness among our bot listening behaviours, the second round of bots (the bots playing Råg i Ryggen's "Queen of Darkness") would be done in a more haphazard way, not following the strict metric logic of the first scenario. Tentative results from the first bot intervention indicated that the regular scheme was unnecessarily rigorous—i.e., too non-human. We desired more varied results, so rather than regularly liking, disliking, or skipping every fifth song, the second round of bots were programmed with a 50 percent probability that they would like, dislike, or skip every third song.

Interactive feedback generated by dislikes, likes, and skips is important to Spotify. Furthermore, likes and dislikes are default user interactions on many contemporary platforms such as Facebook, YouTube, or Twitter. However, in a music streaming environment, the skip is arguably of most significance: "The skip button is now a big part of the overall listening experience" Paul Lamere, the Director of Developer Platform for The Echo Nest, stated in May 2014. Lamere's blog post was centred around how people use "the skip button when listening to music". As an insider, he has access to a vast pool of listener data. For his study, Lamere processed several billions of plays from millions of unique listeners. Basically, his data suggested that when users are engaged with music, they tend to skip more, but when music is in the background, "such as when we are working or relaxing, we skip less . . . The big surprise for me is how often we skip. . . . nearly

every other song that we play” (Lamere 2014).

Our bot interventions cannot really be compared to Lamere’s massive analyses; his scale is different, and as media researchers we had to work outside the data filled black boxes of Spotify and The Echo Nest. Then again, our experiment bears some resemblance with their focus on user activity and music listening behaviours that might alter recommendations. Regarding the functionality of Spotify Radio, one comment on Lamere’s blog post (from reader “Bill”) was particularly striking:

As a radio programmer [your claim] backs up what we’ve known and how we’ve programmed radio stations for years. Because we are BROADCASTERS trying to be MASS APPEAL, this explains why songs are repeated so often (a familiar and popular song has the best chance of keeping most of the listeners) and placement of promos & commercials and when and how often DJ’s talk and what they talk about. Our rating service . . . shows us exactly when people tune out/in and for how long in real time. That’s an eye opener. Our studies also show that people will say they want new music and new music discovery but we watch data that says listeners choose familiarity almost 100% of the time. New, unfamiliar music has a very high skip rate (Bill 2014).

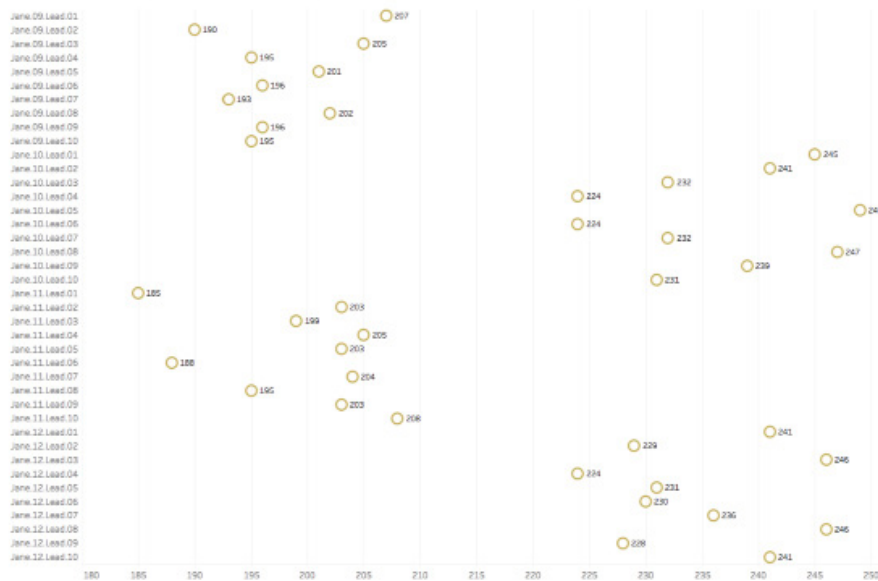
One of the conclusions from our interventions is in fact that artist and to some extent song repetition within Spotify Radio is reminiscent of rotation policy at commercial radio stations. The synergies between record labels and commercial radio naturally have a long and intricate media history. Suffice it to say, constant reiteration and the repeated airing of a limited playlist of songs is a central part of this relation. Songs put on heavy rotation can sometimes be played more than ten times a day so listeners are never far away from the biggest hit at the moment. Still, whether track repetition on Spotify Radio has strict commercial reasons (like at commercial radio broadcasters) remains concealed. Again, as stated, Bas Leijder Havenstroom’s claim—“I believe record labels pay Spotify to have their artists to show up in radio stations”—cannot be objectively answered. The argument if song recommendation within Spotify Radio is commercial and familiarity biased is simply impossible to prove since Spotify does not share data and statistics on the matter. However, our data suggest it might be true.

## Experiment Results

The two Humlab interventions used 160 bot listeners. All bot experiments ran for 24 hours. Chart 1 below displays the amount of plays of 40 Jean Lead bots that all

listened to a radio channel based on Abba’s “Dancing Queen”. Importantly, these 40 bots were divided into ten smaller groups each with the different characteristics of an obedient listener, a liker, a dislike, and a skipper. The setup was similar for the Fred Mercury bots, the Jane Lead bots, and the Carrie Aluminium bots (in the second intervention with a radio channel based on Råg i Ryggen’s “Queen of Darkness”). Taken together, our two bot interventions played a substantial amount of songs on Spotify. In the case of the 40 Jane Lead bots, they listened to more than 7,000 tracks (not including almost 1,700 ads). With some variations, the same goes for the other bot rounds. In the second intervention, with a radio channel based on “Queen of Darkness”, the 40 Carrie Aluminium bots played more than 4,600 tracks. In both of our experiments, 22,624 songs were played. In addition, our bots were programmed to log all 8,367 advertisements featured within Spotify Free.

The substantial amount of plays performed by our 160 bots gives some indications as to how Spotify Radio’s music discovery engine works. Working with bots has its benefits, but it is also problematic. In a former experiment linked to our

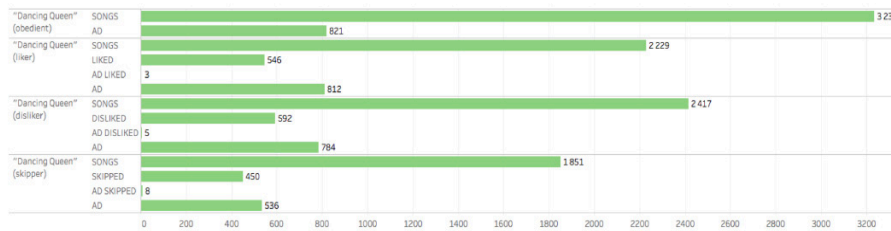


**Chart 1.** Total amounts of songs played by different Jane Lead bots that started a radio channel based on Abba’s “Dancing Queen”, ranging from 185 to 249 tracks.

project we had many problems deploying bots as informants (Snickars & Mähler 2017). In the new experiments with Spotify Radio, our systems were way more robust. They had been tested by frequent and repeated runtimes (and increased knowledge around client logic). Regarding the reliability of our interventions one

should be hesitant, but what can be claimed in the statistics displayed in various charts below indeed supports assertions regarding how Spotify Radio functions. In general, song loops within Spotify Radio tend to be repetitive. Empirically our experiments demonstrate that tracks, especially the same artist, are frequently repeated. As is evident from Chart 1, average sizes of radio loops (during a 24-hour intervention) also vary substantially. Of the 40 Jane Lead bots, each bot listened to approximately 217 tracks (i.e., the average size of these radio loops). Some radio loops played by the Jane Lead bots, however, consisted of around 200 subsequent tracks and others of more than 240 tracks. Basically, the number of songs depended on the length of tracks as well as number (and length) of advertisements.

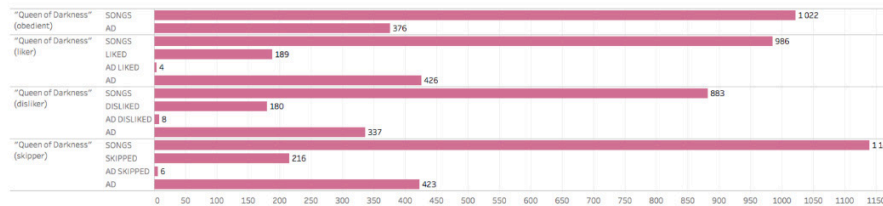
It is important to stress that due to different execution failures, which were computationally identified, all of our bots did not function properly. In both of our interventions, there were 12 failed bot rounds (nine in the first and three in the second). For some unknown reason, this was particularly the case with the 40 Fred Mercury bots, with nine failed play outs – still, the flamboyant lead vocalist



**Chart 2.** Number of songs, advertisements, and interactivity (dislikes, likes, and skips) among approximately 80 bots listening to a radio station based on Abba's "Dancing Queen".

of Queen did at least set the name standard for our subsequent metallic bots. A more severe problem was caused by a software bug, which made the 40 Anna Cobolt bots slightly malfunction, not performing exactly what they had been programmed to do. They basically functioned as they should, but they did not log precisely all interactions in a consistent manner. On the one hand, bugs in the Python script made bots with the dislike characteristic fail to log the very song that was disliked and instead logged the next track. On the other hand, Anna Cobolt bots with the like and dislike characteristic also failed to log the song after an advertisement had appeared. As a result, we realised that statistics from the Anna Cobolt bots deviated slightly from other bot rounds. The Anna Cobolt bots with the other two characteristics (obedient and skipper bots), however, functioned as they should. Hence, in some cases, as is evident in the charts below, results from these specific Anna Cobolt bots have been taken into statistical and comparable accounts.

Nevertheless, because of difficulties in measuring results, we decided to compare statistics within the two bot rounds by excluding the Anna Cobolt bot round as well as the other failed sessions with “invalid bot listening” and to ignore them in our overall statistical analysis. Therefore, in the first round, playing a radio station based on Abba’s “Dancing Queen”, our 111 bots prompted more than 13,000 different forms of content to be played, including songs and advertisements. As is evident from Chart 2, the amount of music tracks varied substantially between the different bot characteristics. The obedient bot listeners played more than 3,200 tracks, while the skipper bots only played 1,850 songs. The amount of plays between the liker and disliker bots were more equal. The same goes for the number of advertisements (around 800, but considerable less for the skipper bots, with only 500). Interestingly, the obedient bot listeners counted about the same number of advertisements even though these bots played many more songs. One general observation from the first intervention is that user interaction (like, dislike, and skip) within Spotify Free seems to have triggered more advertisements at least



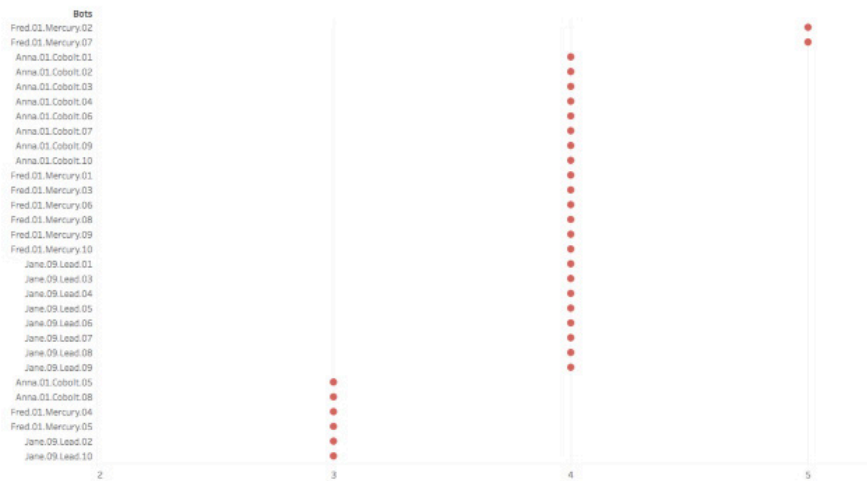
**Chart 3.** Number of songs, advertisements, and interactivity (skips, dislikes, and likes) among approximately 40 bots listening to a radio station based on Råg i Ryggen’s “Queen of Darkness”.

relative to passive bot listeners who did no interaction at all.

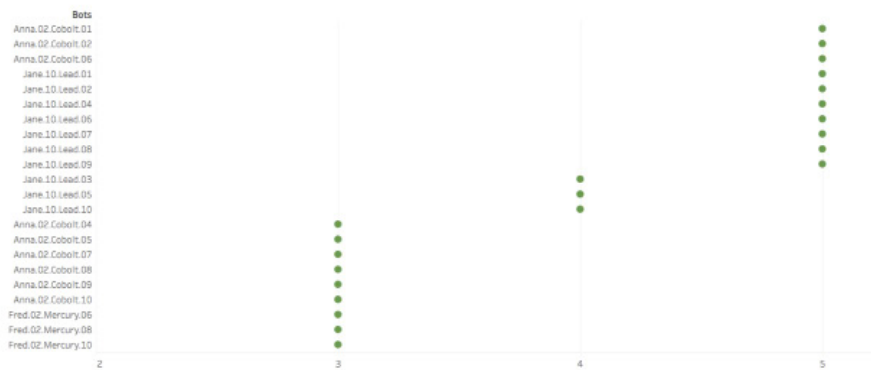
The second round of bots, which started Spotify Radio based on Råg i Ryggen’s “Queen of Darkness”, displayed a similar pattern. Again, the intervention ran for 24 hours, involving 37 bots, and almost 5,600 different forms of content were played, including songs and advertisements. As is evident in Chart 3, the number of advertisements, however, was generally higher. Between 35 to 40 percent of content played were advertisements (compared to the first bot round with around 30 percent advertisements). In relation to the first intervention, it is also striking that the skipper bots, not the obedient bots, were the ones with most songs played. This difference also suggests the difficulty in making general claims about the ways in which the Spotify Radio algorithms work. Notably, however, the second intervention played considerably fewer tracks (approximately 170 songs), compared to the first bot round (approximately 217 songs). The reason for this is mainly that progressive rock (i.e., the similar genre tracks preferred by Spotify’s radio

algorithms) seems to favour extended tracks, so songs were longer and fewer of them were played during a day (and night).

Another way to compare the statistics is to look at how often a song was repeated in the same playlist, i.e., to study repetitiveness in loop patterns. The six charts below depict the amount of repetitions of the same track – first, Abba’s “Dancing Queen” and second, Råg i Ryggen’s “Queen of Darkness” – within radio loops based on the same song. The charts depict the names of our bots with the four different characteristics in recurrent colours: the obedient (red), the liker (blue), the disliker (yellow), and the skipper (green). Again, as stated since the Anna Cobolt bots with the liker and disliker characteristics malfunctioned, these bot characteristics cannot be compared and have been excluded.



**Chart 4.** Amount of repetitions of Abba’s “Dancing Queen” within playlists of obedient bot listeners that started a radio channel based on the same track.



**Chart 5.** Amount of repetitions of Abba’s “Dancing Queen” within playlists of skipper bot listeners that started a radio channel based on the same track.

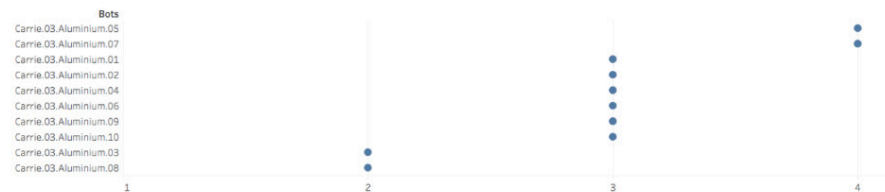


In general, the songs “Dancing Queen” and “Queen of Darkness” repeatedly keep returning among the lists of tracks played by the (unknown) algorithm(s) running Spotify Radio, from twice to five times, and interestingly almost entirely independent of bot characteristics. “Dancing Queen” is repeated more often than “Queen of Darkness”, usually around four times in each radio loop. Statistics for each bot playing “Dancing Queen” reveal that the song kept returning after approximately 50 to 60 tracks. “Queen of Darkness” was repeated less often. Still, on average it kept being repeated around three times in each bot playlist.

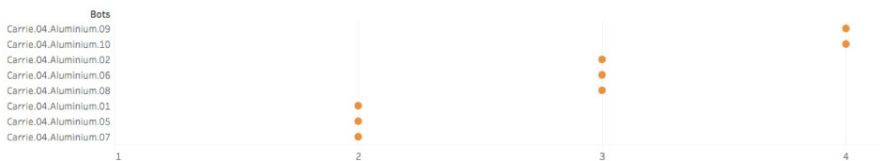
As is apparent from charts 4–9, “Dancing Queen” was repeated more often



**Chart 6.** Amount of repetitions of Råg i Ryggen’s “Queen of Darkness” within playlists of obedient bot listeners that started a radio channel based on the same track.



**Chart 7.** Amount of repetitions of Råg i Ryggen’s “Queen of Darkness” within playlists of liker bot listeners that started a radio channel based on the same track.

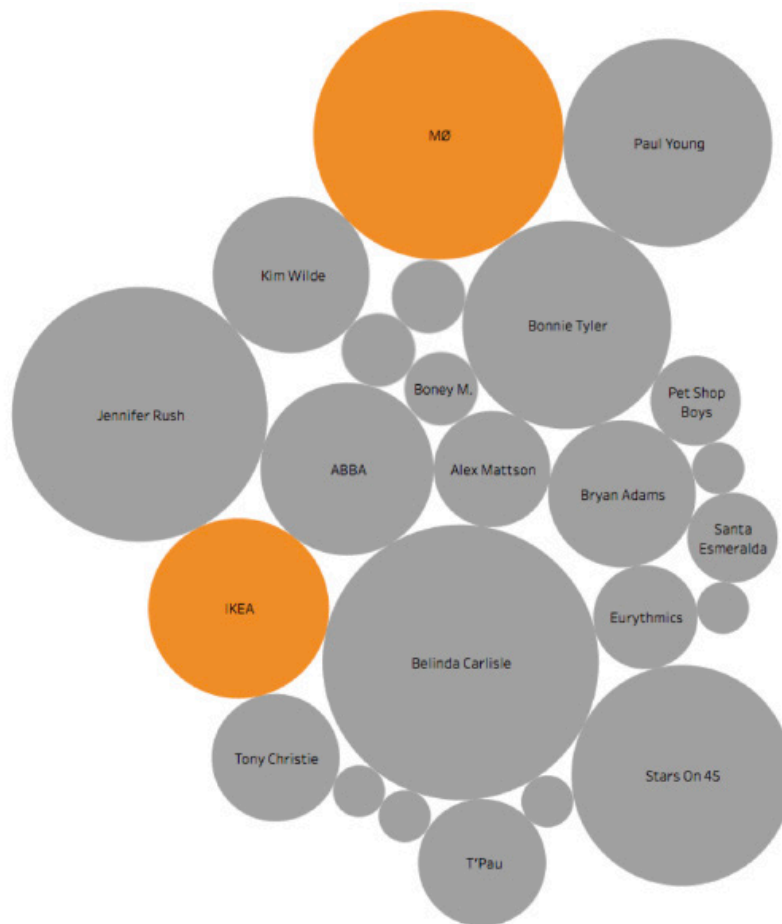


**Chart 8.** Amount of repetitions of Råg i Ryggen’s “Queen of Darkness” within playlists of disliker bot listeners that started a radio channel based on the same track.

than “Queen of Darkness” within our experiments. Another difference between the bot rounds were the general group of artists that Spotify’s radio algorithm(s) generated. In the case of Abba, nearly all recommended artists were strikingly similar, belonging to a homogenous genre of popular hit music from the 1980s (and the late 1970s). The radio stations based on “Queen of Darkness”, however,

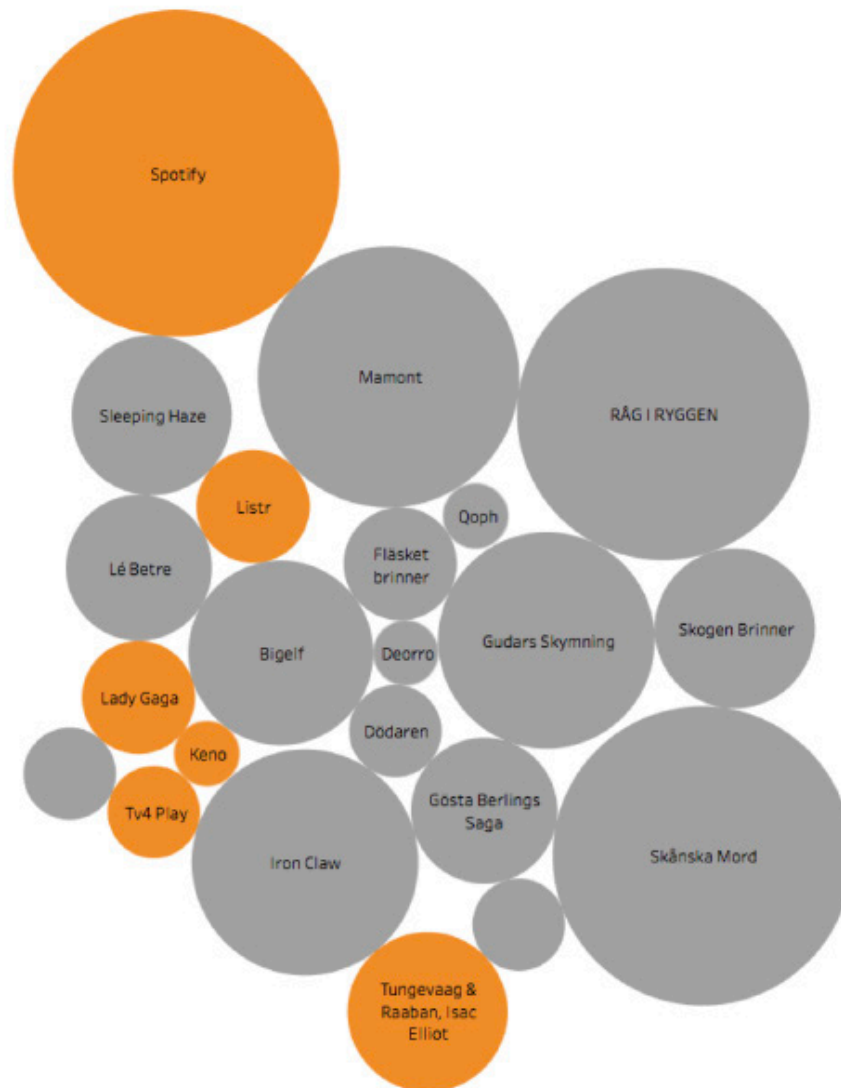


**Chart 9.** Amount of repetitions of Råg i Ryggen's "Queen of Darkness" within playlists of skipper bot listeners that started a radio channel based on the same track.



**Chart 10.** Artists and advertising (orange) in the Abba radio station playlist of liker bot Jane.11.Lead.02.

displayed a much greater variety in terms of artists and songs, and importantly so also from other periods than the 1970s. Songs from the rock band Mamont, for example, were released after 2010, albeit with references to the sounds of 1970s heavy classic rock, progressive, blues, and psychedelic music. The same goes for



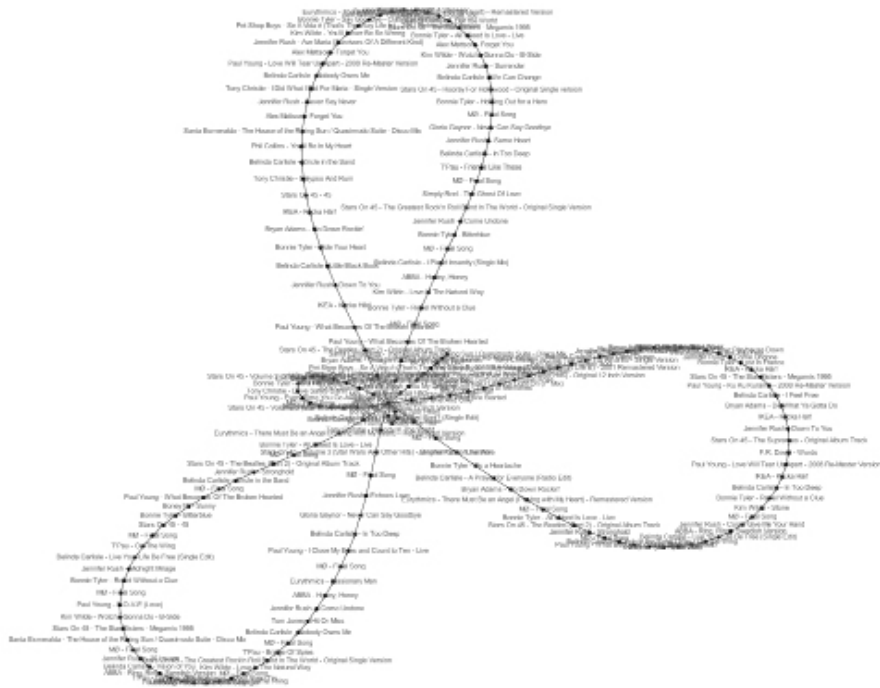
**Chart 11.** Artists and advertising (orange) in the Råg i Ryggen radio station playlist of liker bot Carrie.03.Aluminium.05.

Bigelf, a progressive rock and metal band formed in Los Angeles in 1991, as well as for the contemporary Swedish rock bands Gudars skymning and Skånska Mord (formed in 2006), who both play music influenced by 1970s hard and heavy rock. Temporally, radio stations based on Abba were clearly situated in the late 1970s and early 1980s, while radio stations based on Råg i Ryggen featured a greater variety of music from different periods.

A third dissimilarity between the two bot rounds was the number of advertisements. The two charts above (10 and 11) generically depict advertisement patterns in two radio loops played by two liker bots executed by bots with the most frequent repetitions of “Dancing Queen” and “Queen of Darkness”. As is evident from the size of the circles, advertisement forms a substantial part of content (dis)played within Spotify Free. During the first bot’s playlist, commercials include brands such as Ikea, but also promotions of artists such as Danish singer-songwriter Mø’s “Final Song”. As is evident in Chart 11, more and diversified advertisements were prompted in the second bot playlist, based on (the less popular) “Queen of Darkness”. The playlist featured advertising for brands like Listr, TV4 Play, and Keno, as well as promotions for artists such as Lady Gaga and Tungevaag & Raaban. A general comparison between the two radio loops also makes it evident that Spotify self-promotes its Premium service repeatedly in the playlists of “Queen of Darkness”, with up to 25 ads during the entire intervention, but this heavy promotion was not the case for the Abba playlist.

Chart 12 provides a final way to visualise the quite substantial amount of repetitions apparent within our experiments with Spotify Radio. Chart 12 depicts different song loops that the liker bot Jane.11.Lead.02 listened to after starting a radio station based on “Dancing Queen”; 203 tracks and 42 advertisements were played. The Spotify Radio algorithm prompted “Dancing Queen”, displayed in the middle of the chart although hard to read, to be repeated five times. That is, “Dancing Queen” was played repeatedly as song number 1, 60, 77, 128, 190, and 195, resulting in two minor loops (between song number 60-77 and between 190-195) as well as three major loops. Songs by Abba were played 11 times in the loop, but the most frequent artist constantly reappearing were commercially successful and popular artists from the 1980s such as Belinda Carlisle (28 times), Jennifer Rush (24 times), Paul Young (16 times), Bonnie Tyler (16 times), and Stars On 45 (16 times, plus two promotions).

Even if the song “Dancing Queen” was repeated five times in the loops above, general data from our two bot interventions suggest that loop patterns were not that repetitive, at least not when it came to specific songs. If one takes a closer look at our data around track repetition and how often a song was repeated in the same playlist in the three complete bot playlists, few tracks were played repeatedly. Among the Fred Mercury bots, only six tracks were played five times, and



**Chart 12.** Different song loops as listened to by liker bot Jane.11.Lead.02. The radio channel started with Abba’s “Dancing Queen” (in the middle) with the track being repeated an additional five times (during a 24-hour intervention).

1,990 tracks played one time. For the Jane Lead bots, only two tracks were played six times, but 2,628 tracks were played one time. Within the second intervention, using “Queen of Darkness” to start a radio channel, even less repetitions occurred. Among the Carrie Aluminium bots, 12 tracks were played four times (but no song was repeated more than that). In addition, 2,263 songs were played once.

Repetitions of artists within the algorithms running Spotify Radio is another matter. A closer look at the data behind Chart 12 reveals that the Spotify Radio algorithm picked 20 artists during a 24-hour playout, but only 13 artists were repeated more than twice. Given that Spotify has millions of artists to choose from, it is a relatively low figure. Even without statistically comparing the playlists of all the other bots, artistic patterns within our first bot round was also strikingly similar. Belinda Carlisle, for example, was repeated 32 times by liker bot Jane.11.Lead.07, 28 times by disliker bot Fred.04.Mercury.03, 24 times by skipper bot Fred.02.Mercury.08, and 30 times by obedient bot listener Jane.09.Lead.03. And even if the data for the malfunctioning 39 Anna Cobolt bots cannot exactly be statistically compared, their pattern looks similar as Belinda Carlisle was repeated by all of them.

Given that Spotify has millions of artists to choose from, it appears, at least at first glance, that its radio algorithm(s) did a poor job of playing a variety of artists. Hence, annoyed users at Quora and the Spotify Web Community were partly right: “The terrible radio algorithm repeats the same songs over and over (see [the linked] thread, which has been going for 2+ years)” (tellure 2015) and “need to update the algorithms for Radio, the repetitions are SAD at this point within 1 hour I can easily hear the same song three times” (zaliad 2016). Some songs were indeed recurrent, but it was foremost artist repetition that characterized Spotify Radio in our experiments. Then again, if one bears in mind the ways that Spotify Radio is reminiscent of rotation policies at commercial radio station, repetitions of popular songs and artist can also be perceived as maintaining the status quo. Reiteration is then simply default since regular listeners tend to like similar artists. As stated earlier in this article, such assumptions reveal a normative claim that to work properly the Spotify Radio algorithm should produce apt music recommendations. Perhaps it is sufficient to state that the algorithm(s) behind Spotify Radio are one-sided, since they rarely promote novel artists. Yet, on the other hand, they can also be perceived as commercially biased, although not in a pejorative way but rather from the perspective of mass appeal, as a way of keeping as many listeners as possible tuned to a station.

## Conclusions

One general result from our Humlab bot interventions is that, as part of understanding the inner workings of a central contemporary music delivery platform, it is indeed possible to measure loop patterns on Spotify Radio. This study’s results suggest that bits can be used as a set of concrete methodologies for performing humanist inquiry on big data and black-boxed media services as Spotify, media that today increasingly serve as key delivery mechanisms for cultural materials. The bot logs also made it possible to empirically sustain claims of repetitiveness within Spotify Radio and indeed prove that at least artist iterance is quite striking. The regularity of patterns is in fact prominent, and music loops are definitively not endless. On the contrary, they display a repeated pattern with slight variations depending on which artist a radio station was based on as well as (to lesser extent) bot characteristics. The tracks that we based our radio stations on, for example, kept returning in the bot playlists. If a radio loop started with “Dancing Queen”, after some 50 tracks, it was played again by the Spotify Radio algorithm(s). Bots listening to a radio station based on “Queen of Darkness” displayed a similar tendency, albeit with the difference that the song was not repeated as often as “Dancing Queen” and at shorter intervals (regularly after some 70 tracks or so).

One conclusion to draw from our experiments with Spotify Radio is that si-

milar artists reappear frequently within all bot playlists. It seems that music recommendation algorithms do not take advantage of the archival infinity at Spotify. In short, fans of Belinda Carlisle will be pleased, since a radio station based on “Dancing Queen” will repeatedly play her songs no matter what kinds of interactions are executed. Essentially, the same goes for radio stations based on “Queen of Darkness”, where few artist (yet from different periods) constantly appeared as well.

If Spotify Radio is about personalisation of content, as the company claims, then the recommendation algorithms are a disappointment. This type of algorithmic critique, which the present article can (at least to some extent) empirically verify, taps into contemporary discussions around the ways that machines are altering our taste (for the worse). “Spotify is making you boring”, music journalist Scott Timberg argued in an issue of *Salon* during the summer of 2016. “With all the songs at our fingertips, we’re exposed to very few, thanks to how digital recommendations work”, he stated, which sounds like a description of our interventions. Algorithms, Timberg further wrote, influence listener habits “whether driving your streaming playlists, your Amazon recommendations, or suggestions on iTunes – are about driving you closer and closer to what you already know.” Instead of taking users “toward what you want to listen to, they direct you toward slight variations of what you’re already consuming. (2016)

Timberg’s journalistic speculations, based on personal experiences of massive amounts of music listening, essentially describes most of the results from our bot interventions, at least on a more general level. Radio stations based on our sample songs did result in slight variations of similar music content. One apparent difference, however, between the two bot rounds was that the “Queen of Darkness” radio stations featured a greater variety of music in terms of different periods even if fewer similar artists kept being played. Another difference was that more and diversified advertisements were featured in the Råg i Ryggen bot playlists. In our data, however, we could not really detect specific differences between radio stations based on “Dancing Queen” and “Queen of Darkness”. Our hypothesis that the size and structure of radio loops might depend on music genre as well as popularity was not supported. Indeed, we should have investigated more music genres and more frequently tracked repetition, but artist reiterations indeed were frequent. During both bot interventions, Spotify Radio constantly kept playing more of the same.

An even more troubling result, at least for Spotify, is that radio loops tend to look the same, independent of bot characteristics. Adjusting Spotify Radio through user recommendations such as thumbs up (like), thumbs down (dislike), or skip did, in short, *not* produce differences in results. Although Spotify Radio boasts that “[t]he more you personalise the stations to match your tastes the better

they get”, this is hardly the case, given data from our experiments. At the Spotify Community Web similar comments are also frequent, actually with some users proving their complaints with tangible data. For example, user hahndreas claims that “[g]iving a thumbs-down for songs doesn’t prevent them to be played again just a few songs after. . . . If in a radio-station I’m constantly skipping and thumbing-down slow songs, why doesn’t it move to faster stuff? This should be basic behaviour, shouldn’t it?” (2015). User tamar makes her/his complaint even more directly: “WTF, Spotify?” (2015).

In general, song loops with bots programmed as obedient listeners contained more tracks regarding both “Dancing Queen” and “Queen of Darkness” in our experiments. Otherwise, the data suggest very few real differences between the various bot characteristics. Therefore, an important result from our experiments is that music loop patterns basically look the same, even if bot interaction were performed differently within the Spotify Web client as an obedient bot listener, a liker, a disliker, and a skipper. One reason for this might be that our bots did not have any track record. Following such logic, without previous plays, Spotify’s recommendation engine would have difficulties fine-tuning suggestions and music tastes. However, if one remembers Bas Leijder Havenstroom remark that even if he started a radio station “based on a playlist with many, many artists”, his experience was still “that some (specific) artists keep coming back”. Fine tuning the radio functionality was beyond the bounds of possibility, and the same goes for our interventions. A radio channel based on “Dancing Queen” (nearly) always generated the following artists, regardless of bot characteristics: Bryan Adams, Belinda Carlisle, Toni Christie, Jennifer Rush, Bonnie Tyler, and Paul Young. “Like and unlike buttons are purely decorative”, was user xebec-us’ characterisation of Spotify Radio. xebec continued: “Nothing really gets liked or unliked – you’ll hear the same exact songs with the same frequency as if you were just skipping them” (2015). Even our introduction of randomness among the bot listening behaviours in radio stations based on “Queen of Darkness” did not alter the repeated results generated by the algorithm; more of the same music kept returning.

Another more general conclusion from our experiments is, hence, that the recommendation ability of Spotify Radio is exaggerated. In fact, one might even argue that the claim of musical personalisation and the ability to be recommended an infinity of content to some extent is even untrue. “The more you personalise the better the music gets” should rather be perceived as a mendacious company advertisement used to attract listeners and create commercial interest in the radio functionality (at least when the service was novel). Since complaints were made right after the launch of Spotify Radio, it is likely that the recommendation functionality was flawed from the start. In fact, former Tech Lead, Erik Bernhardsson, said in a lecture in 2013 that “learning from feedback is actually pretty



hard” (Bernhardsson 2013). In addition, the Spotify intern and researcher, Sander Dieleman, also stated that the user feedback that Spotify collects through “thumbs up and thumbs down that users can give to tracks played on radio stations” is hard to measure and make use of and “[t]his type of information is very useful to determine which tracks are similar. Unfortunately, it is also quite noisy” (Dieleman 2014).

Finally, this article concludes that the various forms of public critiques of the inadequate functionality of Spotify Radio are (and were) spot on. In short, it is a service that has not functioned particularly well at least not as a music recommendation system. By and large, however, Spotify seems to have been aware of its malfunctioning radio service and continued to neglect the issue (for different reasons). In an article in *Wired* from May 2016 centred around the latest music discovery releases (Discover Weekly and Release Radar), Edward Newett, the company’s lead software engineer and the person who coded Discover Weekly, problematized the issue: “At the time, I don’t think we were super focused on music discovery in that sense. . . . Spotify had the Discover page, and the artist and song radio, and that seemed good enough” (2014). The interview (together with other accounts) gives the impression that when Spotify Radio was released, it was a premature technology that tried to give listeners good recommendations, but ultimately failed.

Today, the technology seems to have caught up although not in a radio setting. In fact, the radio metaphor within music recommendation systems has been increasingly modified, and recently it has even been altered in favour of other analogies: “We now have more technology than ever before to ensure that if you’re the smallest, strangest musician in the world, doing something that only 20 people in the world will dig, we can now find those 20 people and connect the dots between the artist and listeners”, Matthew Ogle, who oversees Discover Weekly at Spotify, stated in an interview (Pasick 2015). The main ingredient in Discover Weekly is collaborative filtering of user playlists where human selections and groupings of songs form the core of service recommendations. During the last year, Spotify has put way more emphasis on Discover Weekly and recently Release Radar than on its radio functionality. In short, there seems to be a specific tech-musical recommendation narrative, stretching from Spotify Radio (2011) to Discover Weekly (2014) to Release Radar (2016). Unlike Discover Weekly, Release radio’s tracks are brand new and have no listening data. Instead, Spotify relies on a solution that tries to predict who will enjoy a song by analysing the audio signal.

In the end, it seems that traditional radio recommendations appear to be less significant for Spotify, at least in comparison to new types of machinic suggestions. Naturally, feedback data from listener activity and user profiles remain essential for Spotify’s music recommendation systems. Yet, given that the medium

of radio has been around since the 1920s, it is perhaps not surprising that the old radio metaphor (and its cloud counterpart around the celestial jukebox) has finally been superseded. For many years, listeners saw streaming services as a way of getting to know new music; however, as streaming music and not radio became the default listening mode, it is hardly unexpected that the radio metaphor would lose its popularity and consequently be replaced with new computational recommendation formats based on taste profiles, song identification, and digital fingerprints.

**Pelle Snickars** is Professor of Media and Communication Studies, specialising in digital humanities at Umeå university, and is affiliated with the Humlab research centre. His research focuses on the relationship between old and new media, media economy, digitisation of cultural heritage, media history as well as the importance of new technical infrastructures for the humanities.

## References

- Allen Anderson, Paul (2015): "Neo-Muzak and the Business of Mood", *Critical Inquiry* 41:4, 811-840.
- Bernhardsson, Erik (2013): "What data points does Spotify's Radio feature use?" <https://www.quora.com/What-datapoints-does-Spotifys-Radio-feature-use> (Accessed 15/03/17).
- Bernhardsson, Erik (2014): "Music Discovery at Spotify" <http://www.slideshare.net/erikbern/music-recommendations-mlconf-2014> (Accessed 15/03/17).
- Bill (2014) <https://musicmachinery.com/2014/05/02/the-skip/> (Accessed 15/03/17).
- Chandra, Vikram (2013): *Geek Sublime: Writing Fiction, Coding Software*, London: Faber.
- Dieleman, Sander (2014): "Recommending music on Spotify with deep learning" <http://benanne.github.io/2014/08/05/spotify-cnns.html> (Accessed 15/03/17).
- Dieleman, Sander, van den Oord, Aaron & Schrauwen, Benjamin (2013): "Deep content-based music recommendation" <https://biblio.ugent.be/publication/4324554/file/4324567> (Accessed 15/03/17).
- Fixmer, Andy (2012): "Spotify Said Developing Pandora-Like Online Radio Service" <http://www.bloomberg.com/news/articles/2012-04-26/spotify-said-developing-pandora-like-online-radio-service> (Accessed 15/03/17).
- Gehl, Robert W. (2014): *Reverse Engineering Social Media. Software, Culture, and Political Economy in New Media Capitalism*, Philadelphia: Temple University Press.
- Gunnarsson Lorentzen, David (2016): *Following Tweets Around Informetric methodology for the Twittersphere*, Borås: The Swedish School of Library and Information Science.
- Hachman, Mark (2012): "Spotify to Take On Pandora With Radio Service" <http://uk.pc-mag.com/internet-products/64346/news/report-spotify-to-take-on-pandora-with-radio-service> (Accessed 15/03/17).
- hahndreas (2015): "The radio option sucks" <https://community.spotify.com/t5/New-comers-and-Contribution/The-radio-option-sucks/td-p/863585/page/4> (Accessed 15/03/17).
- Friesinger, Günther & Herwig, Jana (2014) (eds.): *The Art of Reverse Engineering*, Bielefeld: Transcript.

- Hietala, Heikki (2014): "Why do my Spotify radio sounds so repetitive?" <https://www.quora.com/Why-do-my-Spotify-radio-sounds-so-repetitive> (Accessed 15/03/17).
- Hu, Yajie. & Ogihara, Mitsunori (2011): "NextOne Player: A Music Recommendation System Based on User Behavior", *12th International Society for Music Information Retrieval Conference*, 103–108.
- Kitchin, Rob (2016): "Thinking critically about and researching algorithms", *Information, Communication & Society*, 24 February, 1-16.
- Kjus, Yngvar (2016): "Musical exploration via streaming services: The Norwegian experience", *Popular Communication*, 14: 3, 127-136.
- Lamere, Paul (2014): "The Skip" <https://musicmachinery.com/2014/05/02/the-skip/> (Accessed 15/03/17).
- lehwarck (2012): "Better Radio Algorithms" <https://community.spotify.com/t5/Closed-Ideas/Better-Radio-Algorithms/idi-p/107485> (Accessed 15/03/17).
- Leijder Havenstroom, Bas (2015): "Why does my Spotify Radio play the same artists over and over for me?" <https://www.quora.com/Why-does-my-Spotify-Radio-play-the-same-artists-over-and-over-for-me> (Accessed 15/03/17).
- Lin, Ning, Tsai, Ping-Chia, Chen, Yu-An & Chen, Homer H. (2014): "Music Recommendation Based on Artist Novelty and Similarity", *Multimedia Signal Processing, IEEE 16th International Workshop*, Jakarta, 1-6.
- Madrigal, Alexis C. (2014): "How Netflix Reverse Engineered Hollywood" *The Atlantic* January 2 <https://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679/> (Accessed 15/03/17).
- Modell, Amanda (2015): "Mapping the Music Genome: Imaginative Geography in Pandora Internet Radio and the Genographic Project", *Media Fields Journal*, 10, 1-13.
- Morris, Jeremy Wade & Powers, Devon (2015): "Control, curation and musical experience in streaming music services" *Creative Industries Journal* 2, 106-122.
- Pasick, Adam (2015): "The magic that makes Spotify's Discover Weekly playlists so damn good" *Quartz* December 21 <http://qz.com/571007/the-magic-that-makes-spotifys-discover-weekly-playlists-so-damn-good/> (Accessed 15/03/17).
- Popper, Ben (2015): "How Spotify's Discover Weekly cracked human curation at internet scale" <http://www.theverge.com/2015/9/30/9416579/spotify-discover-weekly-online-music-curation-interview> (Accessed 15/03/17).
- Quora (2016): "Spotify Radio" <https://www.quora.com/search?q=Spotify+radio> (Accessed 15/03/17).
- Resare, Noa (2013): "How does Spotify radio work" <https://www.quora.com/How-does-Spotify-Radio-work> (Accessed 15/03/17).
- Seaver, Nick (2014a): "On Reverse Engineering. Looking for the cultural work of engineers" <https://medium.com/anthropology-and-algorithms/on-reverse-engineering-d9f5bae87812#9tw36ldh1> (Accessed 15/03/17).
- Seaver, Nick (2014b): "Knowing algorithms" <https://static1.squarespace.com/static/55eb004ee4b0518639d59d9b/t/55ece1bfe4b030b2e8302e1e/1441587647177/seaverMiT8.pdf> (Accessed 15/03/17).
- Seaver, Nick (2016) <http://nickseaver.net/> (Accessed 15/03/17).
- Shao Dingding, Bo & Tao Li, Wang (2009): "Music Recommendation Based on Acoustic Features and User Access Patterns" *IEEE Transactions on Audio, Speech, and Language Processing* 17:8,1602–1611.
- Snickers & Mähler, Roger (2017): "SpotiBot—Turing Testing Spotify" *Digital Humanities Quarterly* forthcoming.
- Spotify Radio (2016) [https://support.spotify.com/us/using\\_spotify/search\\_play/spotify-radio/](https://support.spotify.com/us/using_spotify/search_play/spotify-radio/) (Accessed 15/03/17).
- Tamar (2015): "The radio option sucks" <https://community.spotify.com/t5/Newcomers-and-Contribution/The-radio-option-sucks/td-p/863585/page/4> (Accessed 15/03/17).

- tellure (2015): “‘Only play songs Spotify has never played before’ option for Radio” <https://community.spotify.com/t5/Closed-Ideas/All-Platforms-Radio-quot-Only-play-songs-Spotify-has-never/idi-p/1225168> (Accessed 15/03/17).
- The Echo Nest (2011): “The Echo Nest Powers Spotify Radio” <http://blog.echonest.com/post/14311681173/spotify-radio-the-echo-nest> (Accessed 15/03/17).
- The Echo Nest (2014): “Spotify Acquires Echo Nest” <http://the.echonest.com/press-releases/spotify-acquires-echo-nest/> (Accessed 15/03/17).
- Timberg, Scott (2016): “Spotify is making you boring: When algorithms shape music taste, human curiosity loses” *Salon* June 10 [http://www.salon.com/2016/06/10/spotify\\_is\\_making\\_you\\_boring\\_when\\_algorithms\\_shape\\_music\\_taste\\_human\\_curiosity\\_loses/](http://www.salon.com/2016/06/10/spotify_is_making_you_boring_when_algorithms_shape_music_taste_human_curiosity_loses/) (Accessed 15/03/17).
- Whitman, Brian (2012): “How music recommendation works—and doesn’t work” <http://notes.variogr.am/post/37675885491/how-music-recommendation-works-and-doesnt-work> (Accessed 15/03/17).
- xebec-us (2015): “The radio option sucks” <https://community.spotify.com/t5/New-comers-and-Contribution/The-radio-option-sucks/td-p/863585/page/4> (Accessed 15/03/17).
- zaliad (2016): “Better Radio Algorithms lead to better music” <https://community.spotify.com/t5/Closed-Ideas/All-Platforms-Radio-Better-Radio-Algorithms-lead-to-better-music/idi-p/1333406> (Accessed 15/03/17).



# Studying Ad Targeting with Digital Methods: The Case of Spotify<sup>i</sup>

By Roger Mähler & Patrick Vonderau

## Introduction

Online advertising is a matter of public interest. Ten years ago, many of us would not have cared much about how Facebook, Google or Spotify place ads on their sites, and how they target particular constituencies of buyers or voters. This has changed since 2008, when programmatic advertising was introduced, an automated procedure of ad buying. Programmatic advertising largely lacks human oversight, making advertising an algorithm-driven business. The procedure enabled \$100,000 worth of ads being placed during the 2016 U.S. presidential election by inauthentic accounts that appeared to be affiliated with Russia. It allows Facebook ad buyers to define target groups such as “Jew Hater,” “Second Amendment,” “Hitler did nothing wrong,” or “Nazi party,” which in turn makes it possible to feed such groups with divisive messages. Platforms have taken an active role in spreading misinformation through advertising. They also monitor user behavior on a large scale. Facebook, for instance, obtains detailed dossiers from commercial data brokers about users’ offline lives, and users have limited means to opt out of their data being used (Angwin et al 2016; Madrigal 2017; Meyer 2017).

Spotify, the Swedish music streaming service, has a less controversial reputation. Introducing programmatic ad buying in 2015, however, the company has made no secret of its abilities to collect data on user behavior. In November 2016, Spotify launched a global outdoor ad campaign with ads jokingly showcasing massive aggregate data sets: “Dear 3,749 people who streamed ‘It’s the End of the World as We Know It’ the day of the Brexit vote, hang in there” (Nudd 2016). Spotify does not just collect “an enormous amount of data on what people are listening to, where, and in what context,” as one of its executives stated in public (Terdiman 2015). The company also acts as a private data broker, providing this collection of contextual data to marketers for ad targeting purposes. Spotify offers

“premium brands” its “extraordinarily engaged, first-party-verified audience at scale” (Spotify for Brands 2015). This offer includes “demographic targeting” as well as “content targeting” to reach users with particular habits, mindsets, and tastes that align with a pre-defined target persona. Playlists, “tailored” to specific urban activities (such as “Morning Commute”) and moods (such as “Life Sucks”) are combined with data on genre preferences, age and gender, geography, language, and streaming habits alongside broader interests, lifestyle, and shopping behaviors, fueled by third-party data providers.

Spotify’s desktop interaction design looks very different today from what it did in 2008 or 2012. In the past, user interaction was organized around tracks and search and community-activating features, such as self made playlists. Today, Spotify’s interaction design re-organizes music consumption around behaviors, feelings and moods, channelled through curated playlists and motivational messages that change six times a day. This present situation—where music has become data, and data in turn become contextual material for user profiling at scale—invites us to pause and reflect about the way songs, books, or films are now typically made accessible. How does Spotify’s ‘service’ relate to the European Union’s new General Data Protection Regulation and its provisions on profiling? What is the long-term strategy behind this massive data collection? Facebook is rumored to have an interest in acquiring Spotify, and both companies micro-target users based on their emotional states. This may have repercussions for music and beyond. As psychologist Michal Kosinski and others suggest in a paper entitled, “The Song Is You,” platforms such as Spotify may strategically exploit the link between music choices and personality traits in the near future (Greenberg et al 2016). Kosinski’ model for behavioral prediction is already used by Cambridge Analytica, a firm notorious for “psy-ops” electoral manipulation in support of Brexit and the Trump campaign (Grasseger and Krogerus 2017).

Although ad targeting is in the public interest, there is little reliable information on how it works. Its effects are often under- or overstated. Some see ad targeting as a means to monitor individuals, a view that overlooks that advertising’s “targets” are not individuated human beings but inferred ones. Rather than being “you,” targets are like you: sets of demographic, psychographic, and other data points (audience segments) aggregated via various online sites and bundled together. Others see the ad targeting in Spotify’s free version as largely ineffective. Listeners complain about the lack of proper targeting, noting that some ads did not match basic data sets including age and gender, location of user IP and user language, genre preferences, and listening context. But this wrongly blames on Spotify what in effect are marketer decisions. Most brands do not micro-target their ads but instead opt for broad media reach, depending on overall ad campaign goals and disposable budget.

To study ad targeting, researchers have an inventory of tested methods at their disposal. Media industry researchers often use semi-structured qualitative interviews as direct observation is difficult in media and tech contexts. Our study of Spotify's advertising technology began with such conventional means. We spoke to Spotify, but also interviewed other key stakeholders at business organizations, such as the Internet Advertising Bureau, and companies offering programmatic services, such as Amnet and Starcom. In addition, we retrieved all available information in the public domain related to Spotify's use of such services. Yet as often is the case in media industries research, these kinds of sources often merely provide a work-around because a problem of access to verifiable data persists. Spotify does not reveal with whom the company collaborates in placing ads, for instance. In order to understand who the main actors are in this process, we opted for digital tools to complement data collection. Doing so, we followed the well-established idea to approach "the digital from the inside out, taking up methods that are already embedded in digital infrastructures and practices, and then adapting these to the purposes of social research" (Marres 2017, 84; cf. Hargittai and Sandvig 2015; Rogers 2013).

The remainder of this article gives a brief technical account of how we proceeded. The overall aim of our research was to map the infrastructure of ad serving companies in order to better understand strategies of intermediation in the current platform-based media marketplace. Platforms such as Spotify, Facebook or Google act as market makers in the digital media value chain. Given the degree "multi-layered platformization" (Hölck 2016) is currently developing, we need to have a precise understanding of intermediary strategies and of their key actors (cf. Skeggs and Yuill 2016). Preliminary results of the research are published separately (Vonderau 2017).

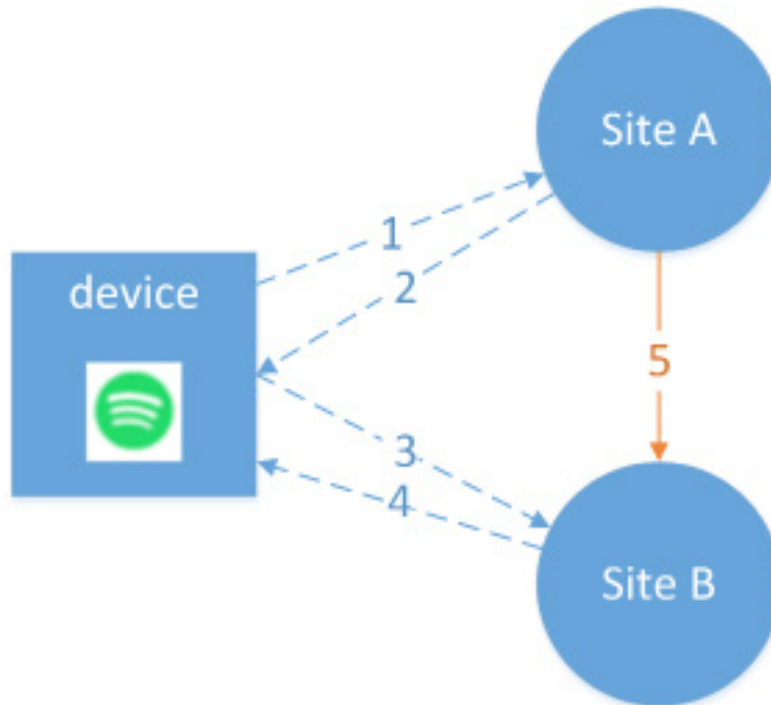
### **Ghostery and Fiddler: Simple Tools for Studying Ad Tech Infrastructure**

Our data work began with opening a small program or plug-in called Ghostery in the browser while being logged in to Spotify's free desktop version. This tool allows us to track or monitor the activities of trackers related to ad programs, analytics, and other functions, and to capture those activities on a trackermap. Ghostery has an extensive, crowd-sourced database of companies that are active within the complex advertising landscape. This allows to identify and chart advertising supply chain vendors.

One limitation, however, is that Ghostery can only be used in conjunction with a Web browser (i.e. the Spotify Web player). In our case, we needed a workflow applicable for all kinds of Spotify clients, not just the Web player, but also desktop applications and Spotify's mobile clients. The purpose was thus to

find an alternative (but similar) workflow that collects Web traffic between the local machine and the remote Web resources such as, for instance, Google’s subsidiary DoubleClick, and ad serving service—all in order to grab (and graph) a snapshot of the underlying ads serving infrastructure. To do this, we needed technical means to intercept and store the communication between actors and companies within the ads landscape.

There are several tools that can be used to capture network data—from generic tools such as WireShark to more specialized tools such as Charles and Fiddler that only capture http and https (encrypted) Web traffic. We used Fiddler, which is a commonly used tool in software development. On Fiddler’s home page



**Figure 1.** The Spotify client requests a resource from site A and receives a response. This response spawns a new request to a resource from Site B.

([www.telerik.com/fiddler](http://www.telerik.com/fiddler)), the tool is described as a “debugging proxy,” which in essence means that the software positions itself as an intermediary layer between the client software and the internet, and in such a way that all client requests, and the associated responses, are routed through this layer. Fiddler can then store and even change or “fiddle with” these messages (hence the name). Fiddler allowed us





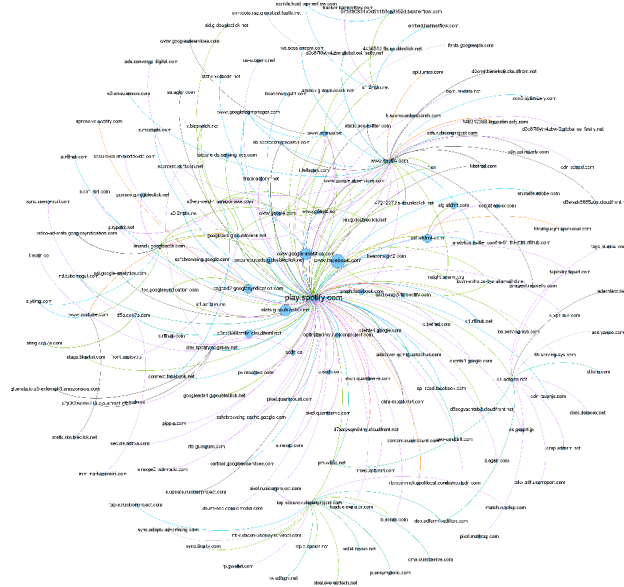
or failed requests.

An hour long session with the Spotify desktop client—logged in using a Facebook account—resulted in no less than 2,391 collected requests (and associated responses). Of these requests, 1,025 were irrelevant for this research, with 1,366 requests remaining. Not all of the latter, however, were ad-related. We got 279 Web documents (html files), 691 images, 209 source code files (JavaScript), 56 text-data files (text, xml or json), 21 files, 54 redirects and 56 of unspecified type (most of them having so called “Not Changed” response code). These requests originate from 17 sites in nine different domains, and targets 71 different hosts in 41 different domains.

The semi-manual workflow used in this part of the experiment started by specifying a usage scenario and context that designated what tracks to play and actions to perform in Spotify. Before we executed our scenario we started the data capture; besides capturing Web traffic, we also recorded the entire desktop session using the open source Open Broadcaster Software. During the session we added comments at specific points to indicate when certain actions started, or when certain updates occurred in the user interface (for instance when a new track started or the leaderboard updates). This made it easier to select and analyze a subset of the traffic that corresponded to specific actions or updates.

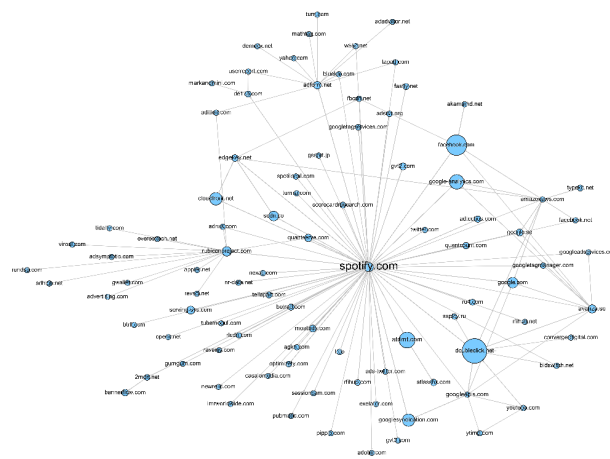
When the usage scenario had ended (i.e. the user was logged out and the recording stopped) the data was exported to Excel for noise elimination and encoding (e.g. content type encoding, domain names, cookie identification, redirects). This was also the step where site names and domain names were translated into actors and companies, with the help of online resources such as the Ghostery database, the Thalamus database, and Cookiepedia. Ideally, this site-to-actor linking should be automated, especially since the online databases also give information of where companies are positioned within the ads ecosystem, which was vital for our investigation. The next step was to use a graph visualization tool such as Gephi to explore and analyze the data—both as a whole, or in part for specific user scenarios. The graph below shows the accumulated data exchange between sites during a 60 minute session. As is evident, a lot of noise still remains to be filtered out especially regarding content types, and parties not related to advertisements. The size of a node is proportional to number of requests to that site.

It is possible to get a cleaner graph if one looks at domains (figure 4.) instead of specific sites. For instance, “4721227.fl.s.doubleclick.net” and “stats.g.doubleclick.net” both belong to the same domain, called “doubleclick.net”. Figure 5, in turn, goes a step further by displaying the companies that operated within each domain. In fact, the graphs become even more interesting if one only selects requests that are related to a specific user action or a system event. For instance, if one



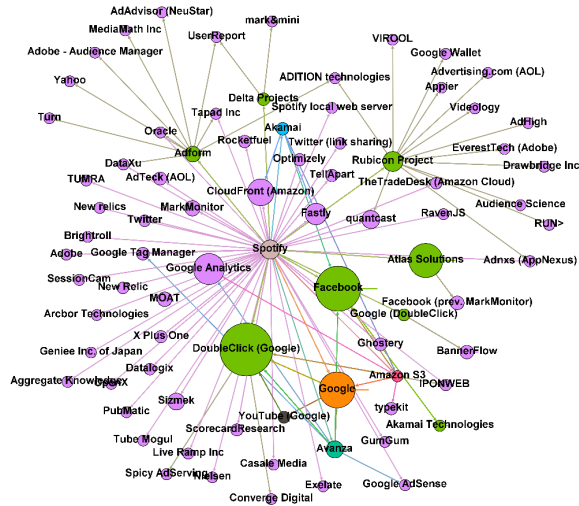
**Figure 4.** Graph of involved internet sites during a 60 minute session.

selects the requests occurring during one single update of Spotify's banner ads, one can create a graph specific for this update, and even schematically show the sequence of how that graph evolved. By using this straightforward, and not overly-complex workflow, one can thus get a number of insights into what is actually going on behind the scenes in the context of a specified Spotify usage scenario.



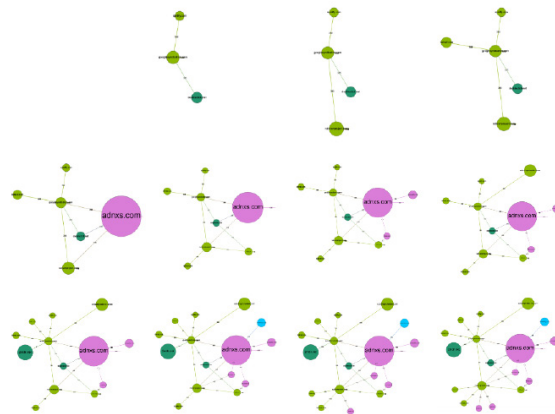
**Figure 5.** Domain graph of involved internet sites during a 60 minute session.

As is apparent, the collected data is rich, which enables one to explore involved parties and messages sent between parties. With this data one could, for example,



**Figure 6.** Graph of involved companies during a 60 minute session. Nodes colored by the Gephi Closeness Centrality algorithm, and edges by type of received content. Redirects and content type JavaScript has been removed from this graph.

relate parties and messages to the portions of the user interface being affected by the interaction. The workflow is rather time-consuming, though, and requires both technical knowledge around tools and Web protocols, as well as domain knowledge of the entire ads ecosystem. It is also helpful if one is knowledgeable about the actors involved—all the way from the advertiser to the targets being exposed to the ads. For this purpose, it is possible to create more automated chains



**Figure 7.** Graphs showing the evolving network of a Spotify Leaderboard as update. The networks are in sequence from left to right and top to bottom, with the final network last.

of tools, especially as there are some public crowd sourced initiatives, an example being the Thalamus database.

## Conclusion

This article has provided a brief description of some digital methods used in studying digital advertising technologies and the key stakeholders involved in ad tech infrastructure. We aimed to balance a more systematical media industries approach (Holt and Perren 2009) with the requirements of object-adequate methods for digital data collection. The aim of this research was not so much to generate representative results or models for other usage scenarios. Rather, it aimed to pinpoint how digital tools can be used in critical research without violating user rights or exposing sensitive company secrets. Ethical guidelines issued by the AOIR-Association of Internet Researchers or the Council for Big Data, Ethics, and Society tend to focus on human subjects in Internet research. Major companies, however, are studied in different ways. Platforms that now act as quasi-monopolies for distributing cultural goods and services need to be open for “audit testing,” and such testing must be made possible despite the often restrictive Terms of Services these platforms define (Sandvig 2017).

**Patrick Vonderau** is Professor at the Department for Media Studies at Stockholm University. His most recent book publication is the co-authored *Spotify Teardown*, forthcoming from MIT Press in 2018. He is a co-editor of *Montage AV*, a co-editor of *Media Industries Journal*, and a co-founder of NECS – European Network for Cinema and Media Studies. E-mail: [patrick.vonderau@ims.su.se](mailto:patrick.vonderau@ims.su.se)

**Roger Mähler** has a degree in computer science and currently holds a position as lead developer at Humlab, Umeå University. Mähler has a long background working as a software architect and systems developer both in academia as well as in the private sector. He participates in several research projects where his main focus is on software development and text analysis. E-mail: [roger.mahler@umu.se](mailto:roger.mahler@umu.se)

## Notes

i Digital tools were used during one week in late August 2016. All these tools are publicly available. The data collection has ended and did not involve user data or sensitive company information. With the public and academic interest in mind, we appreciate Spotify’s forbearance with any trespassings of Terms of Service that our data collection may have involved.

## References

- Angwin, Julia et. al. (2016): 'Facebook Doesn't Tell Users Everything It Really Knows About Them,' ProPublica, December 27. Accessed August 19, 2017. [www.propublica.org](http://www.propublica.org).
- Grasseger, Hannes and Mikael Krogerus (2017): 'The Data that Turned the World Upside Down,' Vice Magazine, January 30. Accessed August 19, 2017. [www.vice.com](http://www.vice.com).
- Greenberg, David M. et al (2016): 'The Song Is You: Attribute Dimensions Reflect Personality,' *Social Psychology and Personality Science* 7(6): 597–605.
- Hölck, Katharina (2016): *Beyond the Single Platform: An Assessment of the Functioning and Regulatory Challenges of Multi-Layered Platform Systems in the Media and Communications Sector*. Brussels: Vrije Universiteit Brussel.
- Holt, Jennifer and Alisa Perren (eds.) (2009): *Media Industries: History, Theory, and Method*. Chichester: Wiley-Blackwell.
- Madriral, Alexis C. (2017): 'Google and Facebook Have Failed Us,' *The Atlantic*, October 2. Accessed October 2, 2017. [www.theatlantic.com](http://www.theatlantic.com).
- Marres, Noortje (2017): *Digital Sociology. The Reinvention of Social Research*. Cambridge: Polity Press.
- Meyer, Robinson (2017): 'Could Facebook Have Caught Its "Jew Hater" Ad Targeting?,' *The Atlantic*, September 15. Accessed October 2, 2017. [www.theatlantic.com](http://www.theatlantic.com).
- Nudd, Tim (2016): 'Spotify Crunches User Data in Fun Ways for This New Global Outdoor Ad Campaign,' *Adweek*, November 29, 2016. Accessed October 24, 2016. [www.adweek.com](http://www.adweek.com).
- Rogers, Richard (2013): *Digital Methods*. Cambridge: MIT Press.
- Sandvig, Christian and Eszter Hargittai (eds.) (2015): *Digital Research Confidential: The Secrets of Studying Behavior Online*. Cambridge, MA: The MIT Press.
- Sandvig, Christian (2017): 'Sandvig v. Sessions: Challenge to CFAA Prohibition on Uncovering Racial Discrimination Online,' [www.aclu.org/cases/sandvig-v-sessions-challenge-cfaa-prohibition-uncovering-racial-discrimination-online](http://www.aclu.org/cases/sandvig-v-sessions-challenge-cfaa-prohibition-uncovering-racial-discrimination-online), September 12. Accessed August 19, 2017.
- Skeggs, Beverley and Simon Yuill (2016): 'Capital experimentation with person/a formation: how Facebook's monetization refigures the relationship between property, personhood and protest,' *Information, Communication & Society* 19 (3), pp. 380-396.
- Spotify (2015): *The Spotify for Brands Team, 'Hello private marketplaces. Spotify here.'* 5 November 2015, [brandsnews.spotify.com](http://brandsnews.spotify.com). Accessed August 19, 2017.
- Terdiman, Daniel (2015): 'Spotify Exec: "We Collect An Enormous Amount of Data",' *Venturebeat*, February 24. Accessed August 19, 2017. [www.venturebeat.com](http://www.venturebeat.com).
- Vonderau, Patrick (2017): "The Spotify Effect: Digital Distribution and Financial Growth", *Television and New Media* (forthcoming).