University of Groningen

Musica Analytica

Prey, Robert

Published in:
Networked Music Cultures

DOI:
[10.1057/978-1-137-58290-4_3](https://doi.org/10.1057/978-1-137-58290-4_3)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Prey, R. (2016). Musica Analytica: The Datafication of Listening. In R. Nowak, & A. Whelan (Eds.), *Networked Music Cultures* (pp. 31-48). (Pop Music, Culture and Identity). Palgrave MacMillan.
https://doi.org/10.1057/978-1-137-58290-4_3

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

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

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Musica Analytica: The Datafication of Listening

Robert Prey

In the winter of 1938, Theodore W. Adorno arrived in New York aboard a steamer from Europe. His friend, the Austrian sociologist Paul Lazarsfeld, had successfully secured a position for Adorno on an ambitious, cutting-edge research project funded by the Rockefeller Foundation. The ‘Radio Research Project’ was an attempt to understand how Americans listened to the radio so that they could be more effectively targeted by broadcasters and advertisers. Adorno assumed the role of Chief of the Music Division.

The project’s crowning achievement was ‘Little Annie’. Officially called the Stanton-Lazarsfeld Program Analyzer, Little Annie was a simple electronic device that allowed research subjects listening to a radio show to indicate approval or disapproval with the click of a button. A rudimentary version of ‘like’ or ‘thumb’ buttons on contemporary music streaming services, Little Annie revolutionised how early radio, film and television audiences were measured.

Little Annie did not impress Adorno though. After only two years on the project, the German Marxist quit. Not surprisingly, he was positively horrified by the ‘culture industry’ that grew out of such cybernetic forms of measurement. Simply put, culture, for Adorno, is immeasurable: ‘I

R. Prey (✉)

University of Groningen, Groningen, Netherlands

reflected that culture was simply the condition that precluded a mentality that tried to measure it' (cited in Müller-Doohm, 2005, p. 247).

What would Adorno think of the real-time data harvesting facilitated by contemporary online music streaming services? Taken together, on-demand music streaming services such as Spotify, Apple Music and Deezer, and personalised online radio services like Pandora Internet Radio, are the fastest-growing sector of the global music industry and represent the future of music distribution and consumption in a post-download era (IFPI 2015).¹ What truly distinguishes these services from previous forms of music consumption, however, is the data feedback loop they generate in real time. On contemporary music streaming services all listening time is data-generating time.

The harvesting and analysis of vast troves of listener data facilitates personalisation of the listening experience. The assumption is that the more accurately a streaming service is able to zero in on the tastes of the individual listener, the more time the listener will spend on a service, and the higher the likelihood that they will convert to a paid subscription package. This represents a sea change in how the music industry operates. With listeners drowning in choice, '[w]hat used to be a question of persuasion', writes Eric Harvey (2014), 'has become a problem of prediction'.

Personalisation on contemporary streaming platforms is not only about the music, however. To date, the subscription model remains a relatively niche market, as most listeners stream music for 'free' on ad-supported versions of these platforms.² To increase advertising revenue, streaming services attempt to target ads more precisely at listeners. They do this by drawing correlations between music taste and the demographic/psychographic characteristics of their listeners. Like the personalization of music, the targeting of ads relies on the harvesting and analysis of vast troves of data.

In what follows, I will explore what I call the datafication of listening. This transformation in how we experience music has broader social implications that I will touch on at the end of this chapter. I begin, though, by describing how such data is collected and used to personalise the listening and advertising experience at two leading streaming platforms: Spotify and Pandora Internet Radio.

3.1 SPOTIFY: THE ECHO NEST

In early 2015, Apple made headlines when it purchased Semetric, the British startup behind the music analytics service Musicmetric. This move signalled that Apple's hotly anticipated new streaming service was ready

to compete with the streaming heavyweight Spotify. Spotify itself had cemented the centrality of data analytics for the streaming sector the year before when it paid a reported \$58 million for a little-known startup called ‘The Echo Nest’ (Ingham 2015). As ‘the world’s leading music intelligence company’, The Echo Nest is a logical place to begin our look at the datafication of listening.

With a knowledge base of more than a trillion data points, covering about 37 million songs and 3.3 million artists, The Echo Nest treats music taste correlation as a scientific problem. To accomplish this seemingly Sisyphean task, The Echo Nest’s acoustic analysis software processes and classifies music according to multiple aural factors—from its pitch to its tempo to its danceability. ‘The system ingests and analyzes the mp3, working to understand every single event in the song, such as a note in a guitar solo or the way in which two notes are connected,’ explains co-founder and chief technology officer (CTO) Brian Whitman. ‘The average song has about 2000 of these “events” for the system to analyze. It then makes connections between that song and other songs with similar progressions or structures’ (as cited in Darer 2012).

At the same time, The Echo Nest conducts semantic analysis of online conversations about music that take place every day, all over the world—millions of blog posts, music reviews, tweets and social media discussions. The Echo Nest platform compiles key words found in descriptions of the music and its creators, and then links them to other artists and songs that have been described with similar key words and phrases. This data is used to determine song similarities on a more cultural level.

Once the world of music has been mapped, the task then becomes to figure out where each individual listener fits on this map, and their individual movements through music space. To this end, The Echo Nest collects a real-time, dynamic record of the type of music fan you are—your music tastes (artists and songs) and music behaviour (favourites, ratings, skips and bans). This is called your ‘Taste Profile’. Taste Profiles are organised into music segments. Such segments are categorised in numerous ways, including artist- and genre-based segments (for example, listeners who like Beyoncé but also like punk music). Other segments are built from listener behaviour (such as listeners who prefer diversity and discovery).

What has been described so far clearly has practical use for song or artist recommendations, but what about targeting ads? How can figuring out what type of listener you are help Spotify figure out what type of ads to show you? With the vast majority of listeners opting for the ‘free’,

ad-supported version, this has increasingly become a central question facing Spotify and other similar streaming services.

When it launched its ‘Music Audience Understanding’ service in 2013, The Echo Nest chief executive officer, Jim Lucchese, was quoted as saying that the company was going from ‘What’s the next song you want to hear?’ to ‘What ads are you likely to respond to?’ (cited in Hof 2013). Apparently, the two are intimately related. Co-founder Brian Whitman even argues that ‘[m]usic preference can predict more about you than anything else’ (as cited in Vanderbilt 2014).

Of course, the entire business of ad-supported broadcast radio has long been predicated on the idea that music taste is an important proxy for listener understanding, and a tool for listener segmentation. What The Echo Nest has done, however, is to take this idea out of the realm of common sense. As Lucchese puts it, ‘[s]aying you’re reaching hip hop fans doesn’t cut it anymore’ (in Hof 2013).

The Echo Nest attempts to provide a big data solution by utilising predictive modelling to analyse streaming music listening behaviour in order to identify the psychographic characteristics of listeners: personality, values, opinions, attitudes, interests and lifestyles. It accomplishes this by comparing music affinity to already known demographic and lifestyle interests. The predictive models that are developed from these incredibly large data sets are continually being trained and validated against ‘ground truth’ survey data.

In this way, The Echo Nest claims to be able to identify ‘statistically meaningful relationships between music taste and non-musical information including age, gender and dozens of lifestyle categories’ (The Echo Nest 2014). Lifestyle categories include ‘Gamers’, ‘Foodies’ and ‘Jetsetters’—consumer categories that can be targeted by relevant advertisers. Music streaming services are determined to increase the rates they can charge advertisers for the right to access their listeners. Segmenting listeners according to standard and custom-designed lifestyle categories is one way to accomplish this. With the help of The Echo Nest’s analytics, Spotify can create dynamic music segments organised by musical genres and behaviour, ad-targeting segments that include demographics and lifestyles, and custom segments based on music listening activity.

One of the problems facing the music streaming sector is that the bulk of ad-supported accounts on services like Spotify, Pandora and Deezer are inactive.³ This means that only a small minority of listeners contributes advertising value to such streaming services. These are the ‘committed,

engaged’ listeners—referred to by The Echo Nest as “high-value” listeners’—listeners who must be distinguished from the “low-value” tourists who disappear from sight’ (The Echo Nest 2013).

By developing a method to identify and represent which category each individual listener fits into, The Echo Nest claims that it can then help music streaming services ‘execute strategies focused specifically on acquiring, engaging, and maximizing the ARPU [average revenue per user] of the likely high-value users’ (ibid.). Essentially this is accomplished in two ways. First, the music artists who correlate with high-value users, and those artists correlated with users less likely to stay with the service, are identified.⁴ A particular user’s approximate future value can then be identified by the artists to whom he or she listens.⁵ Second, the patterns of musical behaviour that best predict high-value listeners are isolated. These patterns are organised into five categories that describe listener behaviour. According to a company white paper, these five categories—or ‘Taste Profile Attributes’ (TPA)—are:

- ‘Adventurousness’: how open the listener is to music outside their ‘musical comfort zone’.
- ‘Diversity’: how varied the listener’s preferred styles and genres are.
- ‘Freshness’: the listener’s relative preference for new and recent artists vs older music.
- ‘Locality’: the relative spread, worldwide, of where the listener’s preferred artists come from.
- ‘Mainstreamness’: the listener’s affinity for well-known artists vs obscure artists. (from The Echo Nest 2013, p. 4)⁶

The claim made by The Echo Nest is that, on the basis of such musical data, it can help a music service:

- Predict each user’s future value to the service.
- Identify musical characteristics of high-value listeners to help the service.
- Tailor user experience towards the high-value group.
- Identify psychographic/affinity characteristics of high-value listeners to help the service monetise that group via targeted advertising. (ibid., p. 2)

TPA were developed to help music streaming services figure out who high-value listeners will be early on, so that they can be proactively focused on, instead of those listeners who will contribute little advertising value to a service. This is an example of what communications scholar Joseph Turow (2008, p. 1) describes as ‘marketing discrimination’, whereby ‘marketers increasingly use computer technologies to generate ever-more-carefully defined customer categories—or niches—that tag consumers as desirable or undesirable for their business’.⁷

The Echo Nest (2013, p. 2) recommends that streaming services create a Taste Profile for a new user as soon as they register, as ‘services can predict from just a few data points whether that user is likely to be a future high-value listener or not’. Even before a profile can be built for a new user, The Echo Nest’s mobile software development tools can predict a user’s value by looking at the music in their mobile library and the particular device through which they are accessing the service. As Whitman revealed in an interview, ‘even the fact that you are using an iPhone instead of an Android we know a lot more about your music taste. It’s low level signals that definitely when combined tell you a lot’ (as cited in White 2014).⁸

Once a streaming service identifies and separates ‘high-value’ listeners from ‘low-value’ listeners, the next step is once again to create and identify interest and affinity segments for the high-value listeners. The Echo Nest has developed a set of affinity models to do just this.⁹ The benefit to music streaming services is obvious: brands and advertisers will pay higher ad rates if they can reach—or more accurately, *think* they can reach—high-value listeners. In short, music streaming space is not only horizontally segmented via consumer categories, it is also vertically ordered via hierarchies of listener value and projections of future worth.

3.2 PANDORA INTERNET RADIO: THE MUSIC GENOME PROJECT

Pandora Internet Radio, the leading audio streaming service in the USA, provides us with another interesting case through which to examine the datafication of listening. As with its streaming service competitors, Pandora collects users’ registration data—age, gender and zip code—and combines this data with insight into the time of day, day of the week and the device that Pandora listeners use to access the service. However, Pandora’s ability to dissect music and to draw correlations between it is dependent on its unique contribution to music taxonomy—the ‘Music Genome Project’.¹⁰

Instead of grouping songs by genre (as record shops and radio stations do), by collaborative filtering ('listeners who like this also like ...') or by ratings, Pandora organises music by musical traits, or 'genes'. Such genes could include, for example, the gender of the lead vocalist, the tempo of the chorus, the level of distortion on the electric guitar, the type of background vocals and many more.¹¹ There is no single, one-size-fits-all list of attributes that applies to all types of music. According to the patent application for the Music Genome Project, the number of genes differs widely between musical genres. Rock and pop songs have 150 genes, rap songs have 350, jazz songs have approximately 400, while world and classical music have between 300 and 500 genes (Glaser et al. 2006). For example, since rap music is lyrically driven, it requires a greater list of subsets of genes within the category of lyrics (rhyme schemes, degree of profanity etc.).

The construction of a music genome is incredibly labour intensive, as it requires intricate analysis by Pandora employees in a process that takes 20–30 minutes per 4-minute song. Pandora does not use machine listening or other forms of automated data extraction. Each attribute, or gene, in a particular song is manually assigned a number between zero and five, in half-integer increments. Such an approach to music analysis is thus very different from The Echo Nest's computer-generated models.

Once all the attributes for a particular song have been entered, the song is placed topographically within a set of other similar songs using a distance function.¹² When a listener chooses a song to start a station, behind the scenes Pandora quickly locates the Genome analysis for that song.¹³ An algorithm then compares the song to the genetic makeup of every song in Pandora's database in order to identify songs that contain similar traits. As stated in the Music Genome Project's patent application, '[t]he matching engine effectively calculates the distance between a source song and the other songs in the database and then sorts the results to yield an adjustable number of closest matches' (ibid.). Those songs deemed 'closest' constitute the new 'station'.

Once the algorithmically chosen song begins to play, the listener is able to give the song a thumbs up if he or she likes it, or a thumbs down if he or she does not. This feedback instantly changes the station's playlist. Thumbs down and that song will not play on that particular station again. Thumbs up and the listener will hear that song, and other songs like it, more frequently. Behind the scenes, this feedback affects the particular weight given to some genes over others. 'By raising the weights of genes

that are important to the individual and reducing the weights of those that are not, the matching process can be made to improve with each use' (ibid.). In other words, through continuous feedback, Pandora learns more and more about each listener's music taste, resulting in a progressively more personalised station.

Every single interaction affects the next song that is delivered to the listener. However, these interactions all carry a different weight. For example, a song skip is considered a less important indication of dissatisfaction than a thumbs down. In an interview with *Forbes*, Pandora's former chief scientist Eric Bieschke explains:

You can skip a song because, while you like it, you have simply heard it too much, or you're just not in the mood for it at that moment. And, if you're a new user, you may be looking for something very specific ... or you may just be curious about how our song selection process works. So, with a skip we have to divine its intent which is not easy. (as cited in Diallo 2013)

However, this does not mean that the skip is unimportant as a data point. 'Rare signals (user interactions) are very valuable and have big implications,' Bieschke points out. 'So if you're a longtime user who never hits the skip button and then you suddenly do it, we pay close attention to that and respond accordingly' (ibid.). Bieschke further reveals that for a brand new user, as few as three user interactions can give Pandora most of what it needs to know to deliver the best results.

3.3 DATA-DRIVEN ADVERTISING ON PANDORA

Pandora has discovered 'that the world of playing the perfect music to people and the world of playing perfect advertising to them are strikingly similar' (Singer 2014). Advertising typically accounts for over 80% of Pandora's yearly revenues, as only a minority of Pandora listeners choose to pay \$36 annually in order to subscribe to the service's ad-free version (Pandora Internet Radio 2015). This makes Pandora much more dependent on advertising than Spotify.

In a New York Times article that explores Pandora's ability to target its listeners with the right ads at the right moment, the reader was asked to imagine the following scenario:

Consider someone who's in an adventurous musical mood on a weekend afternoon ... [T]his listener may be more likely to click on an ad for, say,

adventure travel in Costa Rica than a person in an office on a Monday morning listening to familiar tunes. And that person at the office ... may be more inclined to respond to a more conservative travel ad for a restaurant-and-museum tour of Paris. (Singer 2014)

By measuring the frequency of ad clicks and other methods, Pandora is continuously testing hypotheses such as this.

Pandora has also been digging into its first-party data in order to segment listeners into categories that online advertisers are accustomed to.¹⁴ In late 2013, the service began promoting to ad agencies its first two audience segments—Hispanic listeners and a subsegment for Spanish-speaking listeners.¹⁵ In order to create these two categories, Pandora referred to US census data to locate zip codes with a high percentage of Hispanic or Spanish-speaking residents. Pandora then cross-referenced this information with its user registration data to figure out which listeners fit into this ethnic category. For example, Pandora may have a pool of two million listeners that it knows are Hispanic. By then looking closely at the radio stations listened to by this sample, it can locate other Hispanic listeners across its user base of 250 million-plus listeners (see Joe 2014). Pandora's director of product management Jack Krawczyk has claimed that the company's Hispanic and Spanish-speaking listener segments have proved to be 'at least 10% more accurate than those created based on third-party data' (as cited in Peterson 2013).

Pandora has been steadily rolling out new ad-targeting segments. According to Krawczyk, its goal is to add two new segments every four to six weeks. After the initial success with segmenting Hispanic and Spanish-speaking listeners, the plan is to segment listeners with high household incomes, identified via a combination of census data, registered user information and music taste. Interestingly, Pandora has learnt that listeners who exhibit more eclectic music tastes tend also to have higher incomes than the average listener. According to the anthropologist Nick Seaver, Pandora applies a diversity metric to rank the range of a particular listener's music taste. 'A higher diversity score', remarks Seaver, 'should indicate a higher social status, which means that these listeners can have more expensive ads sold against them' (as cited in Harvey 2014).

Relatively simple correlations between age and music listening behaviour can also provide highly informative insights, according to Pandora. For advertisers looking to reach the parents of young children, for example, Pandora isolates listeners of a certain age and then further segments

those who have listened to children's music. Even the child's age can be estimated by factoring in how long it has been since they last listened to a children's station. Finally, survey data has also been effective in identifying useful correlations. According to Krawczyk, tens of thousands of users were surveyed, and it was discovered that men between the ages of 18 and 34 years who listen to electronic music were the best targets for video game ads (as cited in Peterson [2013](#)).

3.4 POLITICAL AD TARGETING

It is clear that music streaming services have made much of the assumption that our taste in music reveals who we really are. It should therefore not be a surprise to learn that the relationship between music taste and political values is a hot topic for music data analysis. While The Echo Nest has modelled the relationship between music taste and political affiliation on its blog (see Whitman [n.d.](#)), Pandora Internet Radio has taken this one step further. In early 2014, Pandora launched a new ad service aimed at helping political advertisers target Democrat or Republican voters.¹⁶

When registering for a Pandora account you are not required to divulge your politics. As a result, Pandora must determine a listener's political preferences through other means. The service first looks at election result data for a listener's zip code, determining whether a listener's county leans Republican or Democrat. While Pandora has permitted political advertisers to target users based on their zip code since 2011, it has more recently begun supplementing this data with information about users' music tastes and other attributes to create a more valuable profile for advertisers. For example, if you are a staunch country music lover but live in an electoral district that leans Democrat, Pandora claims to be able to predict—with between 75 and 80% accuracy—that you likely vote Republican (see Dwoskin [2014](#)).

It is perhaps no surprise that country music fans lean Republican whereas jazz, reggae and electronic music fans tend to vote for Democrats. However, genre taste can only reveal so much. Pandora claims that it can more precisely locate listeners on the political spectrum by looking at how diverse their music tastes are and the type of music they prefer within that range. As Tom Conrad, former CTO and executive vice-president of product at Pandora explains:

it's not just what genres you like but how much you prefer them within your listening spectrum ... A preference for jazz is only a strong indicator

if it takes up around 35% of what you listen to each month. If you have a jazz station but only listen to it 5% of the time, it isn't a good indicator. (as cited in Joe 2014)

3.5 CONCLUSION

The datafication of listening is a remarkably recent development with transformative implications for marketers, the music industry and music fans. Since the invention of the phonograph in 1877, the individual act of listening to recorded music has been largely shrouded in mist, hidden from the prying eyes of marketers and the music industry.¹⁷ What people listened to, how often they listened to it, when and where it was listened to, were always at best a guess. Even after Nielsen began employing the SoundScan media measurement system in 1992, the music–data feedback loop did not extend beyond the record store checkout counter.¹⁸ What became of an album was unknown once it left the record shop. Perhaps it became the soundtrack for a teenage summer. Or maybe it was purchased as an ill-advised gift, never to be listened to again. The fog that blanketed the radio audience was almost as impenetrable. It has always been very difficult to measure terrestrial radio listenership. Radio ratings are thus, at best, estimates. To borrow Philip Napoli's (2003) useful distinction, the *measured* radio audience is markedly distinct from the *actual* audience.

With online music streaming, however, all listening time has become data-generating time. As Paul Lamere of The Echo Nest discloses, '(e)very time a listener adjusts the volume ... every time they skip a song, every time they search for an artist, or whenever they abandon a listening session, they are telling us a little bit about their music taste' (Lamere 2014). Every signal feeds algorithms that work towards building a profile. Without knowledge of how these processes take place, we are being classified and categorised, perhaps as a 'Jetsetter', maybe as a 'Gamer', and as either a 'low-value listener' or a 'high-value listener'.

Is our taste in music so easily predicted from the other attributes and characteristics that define us? Do our predilections for particular brands match up with our preferences for certain bands? In a review essay that neatly summarises recent debates in the sociology of music consumption, Nick Prior (2013, p. 189) writes that 'music itself and our encounters with it are far more complex than can be conveyed through the idea of social origins, let alone statistical data sets and genre categorizations'.¹⁹ Prior's

assessment may be music to the ears of readers disturbed by the description of how streaming listeners are being measured, categorised and targeted. Nevertheless, particularly for an ad revenue-dependent streaming service such as Pandora, the point is not so much whether ‘reality’ is being accurately reflected, but whether advertisers (and investors) are sufficiently convinced of the service’s data wizardry.²⁰

Furthermore, data is never simply constituted *by* ‘reality’, it is always at the same time constitutive *of* ‘reality’.²¹ Following Eli Pariser’s (2011) persuasive argument in *The Filter Bubble*, the organisation of taste through increasingly detailed measurement and personalisation appears to be leading us towards an increasingly segmented society. At an individual level, the ads that result from Spotify’s mischaracterisation of you as a ‘Gamer’ may cause mild annoyance, but from a sociological perspective, right or wrong, the result is the same—further segmentation. Algorithms have structural implications because they stack onto and amplify already existing differences. Thus, the categorisation of listeners’ ‘musical identity’ not only *reflects* social divisions (between ‘high-value’ and ‘low-value’ listeners, for example), but *reinforces* and even produces new divisions.

Finally, as with every case of data mining, there is also the issue of ‘function creep’—the slow but steady widening of a system or technology beyond its originally intended purpose. Listener data may sit in Spotify or Pandora’s database, or it may one day migrate out, as previously undetermined uses for correlating music taste with some other aspect of our lives are discovered. For example, data collected for the purpose of recommending music may be found to deliver a reliable predictor of financial solvency, IQ score or relationship status. What if a taste for early ’90s Nu Metal indicates a higher propensity to default on a debt repayment? This example may seem far-fetched, even ridiculous, but there are already scores of companies that correlate data on a bewildering array of indicators in order to assess any applicant’s credit worthiness.²² Discovering unanticipated patterns is precisely what data mining is designed to do. As Mark Andrejevic (2014, pp. 179–180) writes:

Perhaps some combination ... of my tastes in food, my birthplace and the climate in the city where I currently live, groups me with others who are proven susceptible to a particular type of illness. The database can determine this long before anyone figures out why. For many of those involved, the why will be irrelevant—what will be important is the probability of the prediction.

At this point we can only guess who might be interested in knowing more about our music tastes, and for what purposes. Once data is released, though, there is no taking it back.

Data-driven music streaming services are but one example of how we are increasingly generating digital traces as we go about our everyday lives, engaged in everyday activities. As a result, the datafication of listening has potential implications that extend far beyond music or ad personalisation. These are not issues that Adorno could likely have imagined in the early 1940s, when he resigned from his position at the ‘Radio Research Project’. Little Annie, it appears, grew up to bequeath a large extended family of methods; methods that measure what Adorno always insisted was immeasurable.

NOTES

1. See also Chaps. 4 and 10.
2. For example, leading streaming service Spotify has at the time of writing over 30 million paid subscribers, out of a total of about 100 million active users. Even with only a roughly 30% ratio of paying users to total users, this is high compared to most of Spotify’s competitors. Pandora Internet Radio has converted a mere 5% of its ‘free’ listeners to its subscription plan (Pandora Internet Radio 2015). See also Chap. 4.
3. A study conducted by MIDiA Consulting suggests that on ad-supported music services in 2012, between 60 and 80% of accounts were essentially inactive (cited in The Echo Nest 2013, p. 2).
4. In a demonstration for an unidentified ‘partner service’, The Echo Nest determined that the ‘high-value’ artists for this service were Florence + The Machine, Bon Iver, fun., Kid Cudi, Ed Sheeran, The Kooks, Of Monsters and Men, Angus & Julia Stone, The Naked And Famous and The Black Keys. ‘Low-value’ artists for this service were Katy Perry, Lil Wayne, Red Hot Chili Peppers, Train, Benjamin Francis Leftwich, Hans Zimmer, Wale, Rusted Root, Swedish House Mafia and The Wanted (The Echo Nest 2013, p. 3).
5. The Echo Nest also shows music streaming services how to employ this knowledge to attract and retain high-value listeners by, for example, promoting ‘high-value’ artists (artists most closely associ-

ated with high-value listeners) through homepage recommendations and in their marketing.

6. The Echo Nest acknowledges that different services, with different users, will need to determine the relative importance of each these attributes in distinguishing between ‘high-value’ and ‘low-value’ listeners. Thus, the TPA can be tuned to deliver the specific parameters (‘more adventurous’, ‘less mainstream’, ‘more diverse’ etc.). Depending on the service, The Echo Nest can also adjust the relative value of music listening behaviour, for example prioritising purchases over shares for services that also sell MP3s.
7. Marketers commonly use the terms ‘targets’ and ‘waste’ to distinguish between these two sets of consumers.
8. Brian Whitman suggests that the future of listener understanding and segmentation will get deeper into how, when and where people actually interact with music. As he noted at a talk at Microsoft, ‘not just what they skip, ban and recommend, but when? Did they just break up with their girlfriend?’ (as cited in Vanderbilt 2014). Much of this scenario is dependent on smartphones continuing to be the critical tool through which to harvest implicit signals that aid in generating contextual knowledge. ‘If you look where the music industry is going,’ Paul Lamere told *Fast Company*, ‘music in the future will be played almost entirely on people’s phones. And your phone knows a lot about you ...’ (as cited in Brownlee 2014).
9. In a company white paper, The Echo Nest reported on an exercise conducted for one of its partner music streaming services. The task was to identify the interests and affinities of this particular service’s highest-value users. The interests and affinities most closely related to these users were social causes, concerts, alcohol, green/eco, outdoor adventure and fashion and style. This of course is information highly valuable to the service, and to any potential advertiser interested in promoting its products or services (see The Echo Nest 2013, p. 5).
10. In 2015 Pandora purchased the music analytics company Next Big Sound to provide even greater insight into the listening activities of its users. According to Billboard, ‘Pandora also plans to use Next Big Sound to help brands choose artists with which to partner and analyze the results’ (Peoples 2015).
11. For more detail, see Glaser et al. 2006.
12. In mathematics, a distance function is a [function](#) that defines a [distance](#) between elements of a [set](#).

13. Pandora is not able to play a specific song on demand because it is classified as a non-interactive streaming service under the Digital Millennium Copyright Act of 1998.
14. At present Pandora's rich data set is only being used in-house to attract advertisers who desire a more targeted audience than traditional terrestrial radio provides. However, according to an article in the trade journal *Advertising Age*, Pandora is open to the possibility of using its data to target listeners on other sites or apps (see Peterson 2013).
15. The popularity of Pandora among the Hispanic population in the USA no doubt drove this decision. The Hispanic market accounts for 10 million monthly users, or 20% of Pandora's total listening audience. What is more, Hispanic listeners are younger, more mobile and more socially connected than the average Pandora listener (see Heine 2012).
16. The pioneer in creating demographic profiles for political advertisers is of course Facebook. During the 2008 US Presidential election, political firms began to target Facebook users by mining the wealth of data that users reveal about their interests and location. Political ads do not come without a cost, however. In its 2013 Financial Results Conference Call, vice-president of Pandora Dominic Paschel admitted that 'the number of calls to customer service goes up dramatically when we accept political ads'.
17. Certain aggregate listening experiences have generated real-time data in the past. For example, in the heyday of the jukebox, the music tastes of precise locales could be determined thanks to mechanised play meters that were built into the boxes (Harvey 2014).
18. See Chap. 10 in this volume.
19. In many ways, what The Echo Nest and its competitors do could be characterised as the big data update to the type of research that sociologists of music consumption have conducted for years. Particularly since *Distinction*, Bourdieu's (1984) landmark study on cultural consumption, scholars have busied themselves constructing or critiquing models relating music tastes to socio-economic status. Exactly how the whole music/society jigsaw puzzle fits together—or indeed if it actually exists in the first place—has been a matter of intense debate (see Atkinson 2011; DeNora 2000; Peterson and Kern 1996; Prior 2013).
20. Thank you to one reviewer for pointing out this distinction.

21. Relatedly, it is still an open question whether, or how, knowledge among streaming listeners that every action is being monitored and assessed might affect listening behaviour itself.
22. For example, the German company Kreditech claims to use up to 8000 data points to assess an application for a loan. This includes information that Kreditech gathers from applicants via Facebook, eBay or Amazon accounts, and even information on precisely how a customer fills out the online application. For example, applicants hurt their chances of receiving a loan if they fill out an application by typing it in all caps, or with no caps (Lobosco 2013). Recently, Facebook announced that it had patented technology that lets lenders examine the credit ratings of a borrower's Facebook friends to determine whether he or she is a good credit risk (Bhattacharya 2015).

BIBLIOGRAPHY

- Andrejevic, M. (2014). Alienation's returns. In C. Fuchs & M. Sandoval (Eds.), *Critique, social media and the information society* (pp. 179–190). New York: Routledge.
- Atkinson, W. (2011). The context and genesis of musical tastes: Omnivorousness debunked, Bourdieu buttressed. *Poetics*, 39, 169–186.
- Bhattacharya, A. (2015, August 4). Facebook patent: Your friends could help you get a loan. Retrieved August 30, 2015, from <http://money.cnn.com/2015/08/04/technology/facebook-loan-patent/>
- Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Cambridge, MA: Harvard University Press.
- Brownlee, J. (2014, February 5). The next big thing in music? Apps that read your mind. Retrieved April 1, 2015, from <http://www.fastcodesign.com/3025991/the-next-big-thing-in-music-apps-that-read-your-mind>
- Darer, M. (2012, September 21). The Echo Nest: Redefining The Internet music experience. Retrieved May 19, 2015, from <http://dyn.com/blog/dyn-dns-client-the-echo-nest-internet-music-streaming-spotify-pandora-online/>
- DeNora, T. (2000). *Music and everyday life*. Cambridge: Cambridge University Press.
- Diallo, A. (2013, October 6). Pandora radio's dominance built on big data edge. Retrieved May 22, 2015, from <http://www.forbes.com/sites/amadouidiallo/2013/10/06/pandora-radios-dominance-built-on-big-data-edge/>
- Dwoskin, E. (2014, February 13). Pandora thinks it knows if you are a republican. Retrieved May 22, 2015, from <http://online.wsj.com/news/articles/SB10001424052702304315004579381393567130078>

- Glaser, W. T., et al. (2006, February 21). Consumer item matching method and system. U.S. Patent No. 7,003,515. Retrieved from <http://www.google.com/patents/US7003515?dq=7,003,515>
- Harvey, E. (2014, April 16). Cover story: Station to station: The past, present, and future of streaming music. Retrieved March 14, 2015, from <http://pitchfork.com/features/cover-story/reader/streaming/>
- Heine, C. (2012, October 5). Why Pandora Ads target Hispanics and Spotify's don't. Retrieved May 22, 2015, from <http://www.adweek.com/news/technology/why-pandora-ads-target-hispanics-and-spotifys-dont-144284>
- Hof, R. (2013, November 21). Like these songs? Then maybe you'll like this ad. Retrieved May 19, 2015, from <http://www.forbes.com/sites/roberthof/2013/11/21/like-these-songs-then-maybe-youll-like-this-ad/>
- IFPI, Digital Music Report. 2015. Retrieved February 2, 2016, from <http://www.ifpi.org/downloads/Digital-Music-Report-2015.pdf>
- Ingham, T. (2015, May 10). Turns out Spotify acquired the Echo Nest for just €50m—*Music business worldwide*. Retrieved August 1, 2015.
- Joe, R. (2014, January 7). CES 2014: Why Pandora is happy that cookies are disappearing—AdExchanger. Retrieved May 22, 2015, from <http://www.adexchanger.com/digital-marketing-2/ces-2014-why-pandora-is-happy-that-cookies-are-disappearing/>
- Lamere, P. (2014, January 14). The zero button music player. *Music Machinery*. Retrieved April 1, 2015, from <http://musicmachinery.com/tag/zero-ui/>
- Lobosco, K. (2013, August 26). Facebook friends could change your credit score. Retrieved August 30, 2015, from <http://money.cnn.com/2013/08/26/technology/social/facebook-credit-score/index.html>
- Müller-Doohm, S. (2005). *Adorno: A biography*. Malden, MA: Polity Press.
- Napoli, P. M. (2003). *Audience economics. Media institutions and the audience*. New York: Columbia University Press.
- Pandora Internet Radio. (2015). Annual report 2015. Retrieved from <http://investor.pandora.com/phoenix.zhtml?c=227956&p=proxy>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. UK: Penguin.
- Peoples, G. (2015, May 19). Pandora acquires next big sound | *billboard*. Retrieved February 3, 2016, from <http://www.billboard.com/articles/business/6568814/pandora-acquires-next-big-sound>
- Peterson, T. (2013, November 14). Pandora looks past the tracking cookie by mining user data. Retrieved May 22, 2015, from <http://adage.com/article/digital/pandora-past-tracking-cookie-mining-user-data/245269/>
- Peterson, R. A., & Kern, R. M. (1996). Changing highbrow taste: From snob to omnivore. *American Sociological Review*, 61, 900–907.
- Prior, N. (2013). Bourdieu and the sociology of music consumption: A critical assessment of recent developments. *Sociology Compass*, 7(3), 181–193.

- Singer, N. (2014, January 4). Listen to Pandora, and it listens back. Retrieved May 21, 2015, from http://www.nytimes.com/2014/01/05/technology/pandora-mines-users-data-to-better-target-ads.html?_r=2&
- The Echo Nest. (2013, August 12). White paper: How music services can acquire, engage, and monetize high-value listeners. Retrieved May 19, 2015, from <http://musicalidentity.echonest.com/post/58057887326/engagement-white-paper>
- The Echo Nest. (2014, January 27). Music audience understanding. Retrieved May 19, 2015, from http://corpcdn.echonest.com.s3.amazonaws.com/filer_private/2014/01/27/audience_understanding_finalx.pdf
- Turow, J. (2008). *Niche envy: Marketing discrimination in the digital age*. Cambridge, MA: MIT Press Books.
- Vanderbilt, T. (2014, February 17). Echo Nest knows your music, your voting choice (Wired UK). Retrieved May 19, 2015, from <http://www.wired.co.uk/magazine/archive/2014/02/features/echo-nest/viewall>
- White, E. (2014, March 26). The Echo Nest CTO Brian Whitman on Spotify Deal, man vs. machine, why Pandora ‘Freaks’ him out (Q&A). Retrieved May 14, 2015, from <http://www.billboard.com/biz/articles/news/digital-and-mobile/5944950/the-echo-nest-cto-brian-whitman-on-spotify-deal-man-vs>
- Whitman, B. (n.d.). How well does music predict your politics? Retrieved May 19, 2015, from <http://notes.variogr.am/post/26869688460/how-well-does-music-predict-your-politics#diverse>