I N S Δ M
JOURNAL OF CONTEMPORARY MUSIC, ART AND TECHNOLOGY

What Does Music Mean to Spotify? An Essay on Musical Significance in the Era of Digital Curation

Asher Tobin Chodos

INSAM Journal of Contemporary Music, Art and Technology
No. 2, Vol. I, July 2019, pp. 36–64.

UDC 78:004.8

Asher Tobin Chodos
University of California San Diego
Department of Music
La Jolla, California, United States

WHAT DOES MUSIC MEAN TO SPOTIFY? AN ESSAY ON MUSICAL SIGNIFICANCE IN THE ERA OF DIGITAL CURATION

Abstract: The growing field of "critical algorithm studies" often addresses the cultural consequences of machine learning, but it has ignored music. The result is that we inhabit a musical culture intimately bound up with various forms of algorithmic mediation, personalization, and "surveillance capitalism" that has largely escaped critical attention. But the issue of algorithmic mediation in music should matter to us, if music matters to us at all. This article lays the groundwork for such critical attention by looking at one major musical application of machine learning: Spotify's automated music recommendation system. In particular, it takes for granted that any musical recommendation – whether made by a person or an algorithm – must necessarily imply a tacit theory of musical meaning. In the case of Spotify, we can make certain claims about that theory, but there are also limits to what we can know about it. Both things – the deductions and the limitations – prove valuable for a critique of automated music curation in general.

Keywords: music information retrieval, music recommendation, machine learning, music semantics, meaning, Spotify, digital culture

One overlooked feature of Spotify's software is that its user experience tends not to discriminate among traditional musical types. Its search box, for example, accepts virtually anything as valid input. Users can enter particular artists, albums, and songs, but they can also enter genres, moods, or other kinds of musical keywords. The resulting recommended materials are equally heterogeneous. Whether we take the "lean in" or "lean back" approach,[1] we are confronted with a mixture of

* Author's contact information: tobinchodos@gmail.com

1 Industry jargon referring to music streaming software that assumes an active ("lean-in") or passive ("lean-back") approach to what music is played.

genres, moods, playlists, or other kinds of "hubs" (Spotify's umbrella term for these variegated musical departure points) as search results. Above all of this diverse suggested material hovers the same inviting "play" button; a hub for "black history is now" is clickable in the exact same way as a "radio" station seeded by Parliament. So is the "artist" Parliament, as is their classic 1978 track, "Flashlight."

This is an important feature of Spotify's software design. This array of clickable options nurtures an impulse for instant gratification and is probably a strategy to maximize user retention.[2] It also means that Spotify is not simply a place to go to hear the music you want, but a place to *learn* about what you want as you make your way through a sea of cute icons that respond to clicks with various kinds of sonic offerings. In other words, Spotify is primarily a music discovery service.[3] As website you visit essentially to explore, Spotify communicates a certain seamless intimacy with the user. Spotify is not a machine that delivers requested goods for a fee; it is an open-ended, benevolent, and exploratory experience in which it is assumed that the data surveilled from your behavior can only enrich your relationship with the program and improve the quality of your recommended content.

It is of course natural for any profit-driven enterprise to want to project this benevolence – and, in a culture of what Shoshana Zuboff calls "surveillance capitalism" (Zuboff 2019), Spotify's practice of surveilling user behavior is an unremarkable example of what has become the dominant business model for tech companies. But it is worth pointing out that music consumption in the digital age was not always this way. Napster and MP3.com, for example, were revolutionary simply because of how much music they made easily available, not for the ingenuity with which they helped users discover new music. Today, since putting 30 million songs within reach is no longer impressive on its own, and because the excess of audio material is harder than ever to make sense of and sort through on your own, music streaming services have, increasingly, needed to become music discovery services.

It is impossible to know exactly how Spotify's music discovery engine works. This is because the system does not work in any one way at any one time for any one user.[4] Additionally, even if it were not subject to constant mutation, the actual algorithm is a carefully guarded trade secret. In spite of the limitations on what we can know about the inner workings of Spotify's music discovery engine, it nevertheless seems straightforwardly true that, no matter how Spotify's recommendations are actually made, the system must in some way be predicated on a notion, explicit or not, of musical meaning. Insofar as any recommendation, by a human or by a machine,

2 See Seaver (2018) for a discussion of how recommendation algorithms increasingly optimize retention over more traditional metrics for recommendation quality.

3 Spotify's gradual transition from a streaming service to a discovery service is discussed in detail below, but it is also evidenced in its recent acquisitions: the Echo Nest (acquired in 2014), Niland, Sonalytic, and MediaTv (all three acquired in 2017) are all companies acquired by Spotify that specialize in automated content curation.

4 See Seaver (2013) for an instructive reminder to remain humble about the "knowability" of recommendation algorithms.

depends on ideas of musical salience and similarity, we can say that Spotify's recommendation service represents a tacit theory of musical meaning.[5] This essay seeks to probe that theory; to make some tentative claims about what its essential contours must be (always acknowledging that the Spotify system is hidden and constantly evolving) while introducing a framework for thinking critically about it.

For good reasons, Spotify's system does not encourage this kind of critical thinking. Questions about the system's implicit theory of musical meaning can only serve to remind users that its theory is just one of many – and therefore not necessarily the best one. The success of the Spotify model depends on communicating that its catalog is both complete and effectively managed – that it has achieved a unique balance of "scale" and "care," to use the words of one of the designers of its recommendation technology (See Whitman 2012). Relativizing the theory of meaning upon which the system depends represents a disconcerting imperfection. If the technology populating my "discover weekly" playlist relies on just one way to construe musical significance, who knows what gems it might be missing, how it might be guiding my consumption habits, manipulating my moods, or shaping my personal identity.

Spotify may not go out of its way to highlight this idea, but the notion that the system is in fact predicated on such a theory can be traced back to one of the first places where Spotify's recommendation technology was laid out: the 2005 doctoral dissertation of Brian Whitman at MIT (Whitman 2005). Although it was published well before Spotify officially launched, Whitman's "Learning the Meaning of Music" introduced the basic outline of the software that would eventually power a hugely successful music intelligence company, The Echo Nest, which Spotify acquired in 2014. As I argue below, some aspects of this technology almost certainly continue to operate in present-day Spotify. And so, Whitman's doctoral dissertation forms a useful, if partial, entry point to Spotify's black box.

As is clear from the title of the dissertation, Whitman proposes this technology while engaging explicitly with the question of musical meaning. He promises that he will be "Learning the Meaning of Music" – but meaning in what sense exactly? To echo Hilary Putnam, one of the few humanistic sources cited by Whitman, what is the "meaning of meaning" in that title (Putnam 1975)? Regardless of how much of this technology is actually used for a given recommendation task by Spotify today, this article contends that a theory of musical meaning gleaned from Whitman's dissertation can be a part of the broader effort to think critically about what music "means" to Spotify. More generally, this can offer a basis for thinking critically about the consequences of the rise of automated curation in music.

This issue is analogous to questions pursued in the discipline of "critical algorithm studies." The idea of embedded bias, for example – the prospect that ostensibly

5 It is important to remember that although it is best known for its "data-driven" approach to music information and its automated personalized recommendations, Spotify actually continues to employ human curators. See Ugwu (n.d.).

objective algorithmic tools will silently encode certain assumptions – is a major theme in this field.[6] As is the related issue of "fairness," which focuses on the real-world consequences of applied machine learning, especially as it concerns social justice and inequality. These issues are clearly explained by Tal Zarsky:

> Any institutional decision that applies or allows algorithms to automatically sort, govern, and decide issues related to human actions makes two crucial assumptions: that human conduct is consistent and that with sufficient data human behavior becomes predictable (Zarsky 2016).

It makes sense that the bulk of the critical attention has, so far, been paid to machine learning applications outside of music. For example, financial institutions have begun to incorporate algorithmic recommendations into their decisions about whether to grant home loans; the question of whether those algorithms will tend to reproduce the structural injustice implicit in their ground truth data is an urgent concern for critics of digital culture. In a similar way, the algorithmic aids used in prison sentencing have been the subject of extensive reporting by, among others, the news organization *Pro Publica*. Machine bias in music, by comparison, feels less urgent. Pierre-Nicholas Schwab, an important figure who writes about fairness in machine learning, even uses music as the paradigmatic case of a place where a lack of fairness does not really matter:

> There is a big difference between a music recommendation service and a news recommendation service. What are the consequences of biased recommendations in a subscription-based service like Spotify? Getting a track recommended that you may not like and will skip. The consequences are small for the consumer (Schwab 2018).

Yet, there are other possible consequences. If we are recommended the same kind of music again and again, what does that do to our musical taste? If playlisting algorithms tend to privilege certain genres over others, do not recommendation engines represent a serious social justice concern? What, in short, are the cultural consequences of a music industry increasingly mediated by the software design decisions of a few large companies?

These considerations, and many others like them, should matter to us if we care at all about what music people are exposed to, and the manner in which our culture relates to that music. In order to investigate these questions, we need to get as strong a sense as possible of how these systems work, and then make informed decisions about how we listen to them. In this article, I look at the notion of musical meaning that is at work in the Spotify algorithm (the currently dominant music

6 See, for example, Powles and Nissenbaum (n.d), which raises the issue of embedded bias while reminding us that seeking to "fix" AI in this way actually represents a concession to its viability and inevitability.

recommendation system). I also ask whether this notion is a good one, and what is lost or gained in the transition to a culture of listening in which automated curation is the norm – a transition which, for better or worse, we are definitely making. The following discussion proceeds in three sections:

* First, I sketch a history of Spotify's development, dispelling some commonly held beliefs about it and showing how it transformed from a streaming company to a discovery company. Here, I argue that automated music recommendation services must necessarily rely on some notion of musical meaning.

* Second, I make a case for why Spotify almost certainly continues to employ some of the techniques Whitman developed in his 2005 dissertation.

* Finally, I attempt to discern Spotify's theory of musical meaning itself. I do this, first, via a close reading of the behavior of the Spotify graphical user interface (GUI) and, second, via Whitman's 2005 dissertation. In the latter case, I argue that the techniques outlined in the dissertation are novel and probably effective, but that there are interesting gray areas where Whitman addresses the question of musical meaning. In the end, I neither condemn nor endorse Spotify's system. Instead, I merely hope to show that a system like Spotify inevitably relies upon a theory of meaning; as users of that system we will benefit from paying close attention to what that theory is.

I. Spotify and the "Curatorial Turn" (2008–2018)

There is a widely held belief that when Spotify was launched in 2008, it was as a response to a music industry imperiled by the growing practice of music piracy.[7] It is true that by the time Daniel Ek and Martin Lorentzon created the startup that would eventually mature into a publicly traded corporation worth more than $20 billion, the recording industry had contracted enormously from its peak at the end of the 20th century.[8] The familiar narrative casts Spotify as a reaction to and, perhaps, a solution for the industry's financial crisis. And indeed, this sometimes seems fair: according to the International Federation of the Phonographic Industry (IFPI), for example, industry revenue in 2018 had recovered to 68.4% of that peak value, largely on the strength of a 45% growth in paid subscription streaming (IFPI 2018). As Spotify is by far the largest paid subscription service, with some 200 million active users today (87 million of whom are paying for subscriptions),[9] Spotify appears to be, from this perspective, an important driver of the industry's recovery, vindicating

7 See for example, the BBC news 2018 article "How Spotify came to be worth billions," (BBC 2018) which casts Spotify as a "response to the growing piracy problem," or Silva (n.d.), or many others that echo this idea.

8 According to Greg Kot, revenues from recorded music in America plunged from their all-time peak of $14.6 billion in 1999 to $12.6 billion in 2002, a decline of 13.7 percent. (Kot 2009, 31)

9 Apple music, though, is gaining on Spotify, with 56 million users as of time of writing (see Yoo 2019).

the altruistic posture the company occasionally projects.[10] The major record labels are frequently castigated for their repeated failures to develop viable systems of electronic distribution in the digital age. Spotify, as a kind of commercial imitation of illegal file sharing, can be seen as the music industry's belated effort to rectify that mistake. Heralded as the "solution to music piracy,"[11] Spotify is thought to restore value to the industry, connecting listeners with the music they want to hear and artists with interested audiences – and all in conformity with US copyright law. So goes, at any rate, the familiar narrative.

This narrative, however, obscures some important facts about Spotify and the relationship between music streaming and the music industry in general. First of all, it ignores the fact that Spotify has yet to turn a profit. In fact, Spotify's annual operating losses have increased sharply every single year, from €98 million in 2013 to €378 million in 2017 (Richter 2018). In 2018 and 2019, Spotify's losses have decreased, but the company remains unprofitable.[12] Although these kinds of consistent losses are not unheard of in the present investor climate,[13] Spotify's financial profile should still give pause to those who want to see it as the music industry's savior.[14] It will, after all, eventually have to turn a profit or fold. Moreover, it is important to note that these losses are not for lack of revenue or a reliable customer base, but instead point to the same old problem the music industry has always faced in the digital age: these losses are due primarily to the licensing costs paid out to the major labels, which represent Spotify's biggest operating expense by far. The fact is that customers are unwilling to pay what they used to pay for music, but major record labels remain committed to intellectual property paradigms from the 20th century, paradigms that only work with 20th century revenue streams. This has been the problem facing music sellers for the last two decades, and Spotify has not solved it. If Spotify is responding to an industry beleaguered by widespread piracy, its response fails in precisely the same way that Napster's did. The difference is that where Napster was bankrupted by aggressive litigation from the Recording Industry Association of American (RIAA), Spotify is kept from turning a profit as it funnels most of its revenue (and shares of its stock) to the major labels – which

10 As it does, for example, in Brian Whitman's lengthy 2012 blog post, "How Music Recommendation Works–and doesn't work" (Whitman 2012), discussed at length below.

11 Marsha Silva, "Spotify, the 'Solution to Music Piracy,' Is Getting Pirated by 2 Million People," published in Digital Music News, at https://www.digitalmusicnews.com/2018/03/26/spotify-piracy-hacked/, accessed May 21, 2019

12 See Spotify's publicly available financial disclosures at https://investors.spotify.com/financials/default.aspx

13 Pandora too posts losses in the hundreds of millions, and in general traditional notions of value have changed radically across the economy. As hedge fund manager David Einhorn puts it, "the market has adopted an alternative paradigm for calculating equity value." (quoted in Kim (2017))

14 Spotify's 2019 press release for investors is jubilant about its first quarter earnings while predicting another loss of €180-340 million. See https://investors.spotify.com/financials/press-release-details/2019/Spotify-Technology-SA-Announces-Financial-Results-for-First-Quarter-2019/default.aspx

constitute the controlling forces of the RIAA.

The fact that this can be said of the streaming industry's biggest player raises important questions about the financial viability of the streaming model itself; if Spotify can't make it work, one wonders, who can? Spotify has over the years shifted between various strategies for earning revenue: early on it looked to advertising, before attempting to monetize its integration with Facebook, and now it sees subscriptions as its principal revenue stream. But it would be more accurate to say that Spotify's true source of revenue has always been venture capital, which it has attracted with extraordinary success, gaining more and more money over the course of 24 funding rounds even in the face of large losses. If Spotify succeeds only in raising venture capital, growing quickly, and collecting potentially monetizable user data, it no more represents a solution for the music industry than Uber or Air B&B – both are companies that have been extraordinarily successful at raising venture capital, but which contain no special insights about the music industry.

This familiar narrative about Spotify, in which it is lumped together with other "disruptive" tech firms, also obscures another important fact: that, although it is marketed as a novel and innovative firm, it is in fact largely owned by the traditional music industry forces. Since Spotify cannot afford a market rate for the licensing fees its service requires, it has been forced to compensate the major labels, in part, with company equity rather than cash. As a result, Peter Tschmuck reports, major labels own as much as 20% of Spotify today (Tschmuck 2017, 179). This fact is perhaps the cause of the widespread concern in the music industry about so-called "playola," a word that refers to the influence major labels supposedly wield over the content of Spotify's curated playlists (not to be confused with the familiar "payola," which denotes a similar practice from radio broadcasting).[15] It is also a possible cause for the often reported homogeneity of Spotify's automated recommendations, an effect which, if authentic, would undermine Spotify's stated aims as a music discovery service.[16] In any case, it is important to remember that, although Spotify is often said to have "disrupted" the industry, it is largely owned by the major record labels, and they are the ones who benefit and receive the majority of its revenue.

The familiar narrative also overstates the relationship between Spotify and the industry as a whole. If we believe that Spotify has the potential to "rescue" the industry from the scourge of piracy, we must believe that it has a marked effect on the market itself. Yet, that may not be true at all. While Pandora has commissioned studies showing that Internet radio has positive effects on music consumption in general, there is little consensus on this point and other scholars have found quite the opposite result. Or, we may simply find that Spotify has no net effect on the music industry whatsoever. Aguiar and Waldfogel, for example, find that while Spotify does displace some lost revenue due to piracy, the new revenue is "roughly

15 For a representative complaint about playola, see, e.g. Peoples (2015).

16 Spotify's app blurb on the Google app store, for example, promises "the right music for every moment" (and, moreover, for every individual user) – not just what the major labels want to promote.

offset by revenue reductions from the sale of permanent downloads" (Aguiar and Waldfogel 2015, 22). Spotify stimulates the market in some ways while depressing it in others, and it seems impossible to know exactly how to gauge its impact on the industry as a whole. Therefore, it is not necessarily reasonable to assume that Spotify has either "rescued" or depressed the market.

Furthermore, it is not even certain that the industry's crisis in 2006 was due to piracy in the first place (the problem to which streaming is often seen as a solution). While it is true that by 2006 revenues had seen a sharp decline from their peak in the 1990s, the golden years the record industry enjoyed in the 1990s should not necessarily be seen as the norm. Instead, some have seen them as anomalous, a period of growth artificially stimulated by the advent of the CD and, therefore, inherently short-lived. Revenues had, in fact, been declining for a long time before the arrival of the CD, which gave the industry a lift largely thanks to its new ability to sell consumers CD versions of music they already owned on vinyl and tape. From this perspective, it is only reasonable to expect that this lift would be temporary – and therefore, perhaps it's inaccurate to blame the downturn on internet piracy and file sharing. The claim that piracy is responsible for the industry's downtown, though repeated constantly by the RIAA and industry insiders, is not necessarily true. As Greg Kot notes,

> It was disingenuous of the industry to blame its slump on file sharing without acknowledging the role played by rising CD prices. The average retail price of CDs had increased more than 19 percent from 1998 to 2002. Peak price was $18.99, with middlemen getting the vast majority of the split (Kot 2009, 42).

If this picture is accurate – if the industry's pains at the turn of the century were a natural regression rather than the result of disruptive new technologies or cultural shifts – then the whole idea of Spotify as the industry's savior, "restoring value" to a business struggling to accommodate new technological paradigms, is an oversimplification. Despite aligning itself with the rhetoric of disruptive innovation popular in the tech industry, in actuality Spotify is probably neither the industry's savior nor its destroyer, and, in many ways, it continues the patterns and promotes the interests of the major record labels who are among its largest shareholders. From a business perspective, Spotify is much less exceptional than it seems.

Meaning and the Curatorial Turn

But even if Spotify may not be the determining factor behind a sudden shift in the music industry, it certainly marks one. How (or whether) the streaming industry is to become self-sustaining remains a mystery; nevertheless, it is hard to imagine a future in which the music industry does not have, at its center, music streaming services. Over the last 11 years, Spotify has evolved from a music streaming company that in many ways inherited the mantle of Napster, Gnutella, and Limewire, merely seeking to provide legal access to a large catalogue of music, to a music discovery company whose most valuable properties are its recommendation engines. In this section, I trace that evolution.

In the only academic history of Spotify, Maria Eriksson et al. (2019) divide its evolution into seven periods (Eriksson et al. 2019, 43-67):

- "Beta" Period (2007–2008). Spotify released to a small circle of personal acquaintances.
- Period A (2008–2009). First public version launched in October 2008 in eight European countries. Spotify removes unlicensed music from its service. Spotify begins to sell advertising and launches ad-free Spotify Premium.
- Period B (late 2009). Global financial crisis eats into advertising revenue and venture capital. Doubts about viability of an ad-supported model leads to increased emphasis on subscription services.
- Period C (2010–2011). Spotify as a platform, emphasis on social features. Linking of Spotify and Facebook, increased practice of data extraction from users. "Related artists" function added. Spotify opens in the US.
- Period D (2011–2012). Valuation reaches $10 billion. Increased "platformization." Competition with Internet radio sites in the US (such as Pandora) leads to increased importance of recommendation and discovery.
- Period E (2013). Spotify begins to address "the abundance of choice" as a primary problem. Solution is no longer primarily social, but algorithmic. Spotify positions itself as a discovery company. Spotify acquires music recommendation company Tunigo (May 2013), which recommends music based on social activities and moods.
- Period F (2013–2015). Spotify dismantles the P2P network, opting instead to use its own servers. Spotify acquires The Echo Nest (2014), an important music information company, for $100 million.
- Period G (2015–2016). In competition with Apple Music, Spotify emphasizes its ability to create musical experiences tailored to each moment. Curation strategy combines the expertise of two acquired companies: Tunigo (expert human curation) and Echo Nest (scalable algorithmic curation). Also acquires Seed Scientific, a data science company. Summer 2015, Spotify introduces various personalized weekly playlists, such as "discover weekly."

As this timeline shows, since its founding, Spotify has nimbly adjusted to shifting market priorities and trends in startup culture, at times making dramatic adjustments to its marketing strategy and business model to accommodate these shifts. Not long after the collapse of Napster, Spotify began as a peer-to-peer sharing service that not only copied parts of Napster's technical architecture, but actually permitted the sharing of unlicensed music. When Spotify launched its first publicly available version in 2008, it removed the unlicensed music, but preserved much of the P2P architecture and kept the disruptive caché of Napster as part of its marketing strategy. After the global financial crisis cast widespread doubt on the viability of advertising for all Internet companies, Spotify recast its free tier as a marketing strategy for its subscription service, which would now become its primary revenue stream. In the wake of Facebook's monumental growth around 2010, Spotify partnered with Facebook and integrated itself into the social network giant.

Among these various adjustments, the most important one for the purposes of this paper is the so-called "curatorial turn:" the shift toward music curation as an important element in Spotify's service. Largely because of its arrival in the USA market in 2012, where it had to compete with Pandora and other Internet radio services, Spotify has increasingly positioned itself as a "music discovery service" rather than simply a music streaming service – and this remains the form Spotify takes today. Even a cursory look at Spotify's service today reveals how central recommendations are to its service. This shift can also be seen by looking at the contrast between two versions of Spotify's homepage, one from 2006 and one from Spotify's "about" section in 2019.

In 2006:

> Spotify gives you the music you want, when you want it.
> Your choice is just a search box or a friendly recommendation away.
> You'll be amazed by the speed and control you have with Spotify.[17]

And in 2019:

> With Spotify, it's easy to find the right music for every moment.
> Choose what you want to listen to, or let Spotify surprise you.
> Soundtrack your life with Spotify.[18]

The difference in tone is subtle but illustrative. In 2006, Spotify is a service that, ultimately, delivers "your choice," even if that choice can be optionally mediated by the service's recommendations (recommendations which, at the time, were probably mostly made by humans rather than machines). The leading line promises "the music you want," clearly prioritizing and emphasizing the volition of the user. This blurb also promises the user "speed and control," two features that an informed, self-directed user might value. It clearly targets a user that takes an active role in her media consumption, using what the industry terms a"lean-in" strategy.

Although it probably holds appeal for aficionados and professionals, this posture eventually became a liability,[19] and Spotify had to adjust. And this meant designing

17 Accessed via the Internet Archive, at https://web.archive.org/web/20061127231638/http://www.spotify.com/, accessed May 20, 2019.

18 Spotify Home Page, "About Us," https://www.spotify.com/us/about-us/contact/, accessed May 20, 2019.

19 In 2011, for example, Billboard published an article in which Spotify was negatively characterized

a software that had something to say about musical quality, about music qua music. In 2019, what matters is no longer the music you want, but the music that is appropriate for "every moment." The value the user might find in having control over the tool is replaced by its power to "soundtrack your life," that is, to find music that matches whatever non-musical activity you happen to be engaged in. This is a notable shift to a "lean-back" approach, a shift which has taken place with respect to the media industry in general over this decade.[20] Interestingly, this shift engenders an adjustment in Spotify's attitude toward music itself; as we lean back, music's value comes to reside primarily in its relationship to things outside of itself. A peculiar feature of the rise of curation is that the value of music is based on how it "goes with" other things rather than what it sounds like (a fact which is discussed at greater length below). This is not a posture Spotify found itself taking before the curatorial turn.

More than the size of the catalogue or the quality of the sound,[21] Spotify's current selling point is its discovery product. And although Spotify does continue to employ human curators (See Ugwu, n.d.), it probably uses more automation than any of its competitors. Spotify's service, then, is not simply to provide customers with access to an enormous database,[22] nor is it exactly to help them find music they like. Instead, what Spotify promises is to help customers find the right music for a given moment, to "soundtrack your life." On the face of it, this slogan makes a pretty bold statement: that the millions of tracks in Spotify's catalogue are "soundtrack" music. It is only made obliquely, so it is easy to miss, but it is a real consequence of the curatorial turn. Here Spotify is part of a broader trend in digital culture. As Peter Wikstrom puts it,

> In a world where information is abundant, people may not be willing to pay a premium for basic access to that information, but they are most likely willing to pay for services which help them navigate through the vast amounts of information (Wikstrom 2013, 7).

Spotify is not unique in its turn toward automated curation, but making that turn engenders certain shifts in its basic attitude toward the meaning of music. One such shift is the subtle creep of the "soundtrack," the idea that music is generally supplemental to other activities and modes of consumption.

as "just a huge database of songs." (cited in Eriksson et al 2019, 59).

20 For work on the rise of curation in general, see, e.g. Silberman (2015) and Gillespie (2011).

21 Even the sonic watermarks imposed by many of Spotify's music industry partners (which are noticeable) seem not to deter customers at all. See Matt Montag's blog (https://www.mattmontag.com/music/universals-audible-watermark) for a useful demonstration of those watermarks. Accessed May 16, 2019.

22 Spotify's 30 million track catalogue, while bigger than those of its competitors, is no longer really its main selling point.

II. Is Spotify Using the Echo Nest?

The rise in demand for curation services was an engineering problem that Spotify approached in more than one way. Following the broader trend of social networking after 2010, Spotify's first solution was, to use the industry's word, "social." In 2010, Spotify received $16 million in venture capital from Sean Parker, the co-founder of Napster. After Napster, Parker had gone on to become the founding president of Facebook. With his investment in Spotify, he earned a spot on its board of directors and ensured that the two companies could integrate their products smoothly. Through the integration of Spotify and Facebook, the social model of music discovery was possible: the listening habits of one's friends could be distilled and transformed into music recommendations. This strategy has the advantage of requiring relatively little engineering, and it is predicated on the intuitively reasonable assumption that people share musical tastes with their social groups. There are a number of ways in which this strategy is not particularly useful, though: first, it will never be a reliable way to expose users to music that is not already popular. Second, like all "context based" recommendation systems, it bears no formal relationship to the musical content itself. Third, it still demands the active engagement of the user, the "lean-in" attitude that Spotify had traditionally envisioned for its customers.

Automated recommendations could potentially address these shortcomings. Facing these issues, as well as competition from American Internet radio stations like Pandora, Spotify began to more aggressively develop its automated recommendation engine in 2012. It began to foreground its recommendation services, adjust its marketing strategy, and, above all, it acquired prominent companies in the music intelligence and recommendation space.

Probably the most important acquisition was The Echo Nest, which Spotify bought in 2014 for $100 million (Lunden 2014). Founded in 2005 by two graduates of the MIT Media Lab, Tristan Jehan and Brian Whitman, the company quickly grew into one of the biggest players in the music recommendation space. Its API powered the music recommendation services of major companies like MTV, Rdio, and Spotify (before the latter bought it). The technology employed by The Echo Nest is described in the academic writing of its founders (especially Whitman's dissertation), and below I will be using those texts to make some deductions about Spotify's current software. But is it reasonable to assume that Spotify is actually still using the technology it acquired in 2014? It is widely known, after all, that Silicon Valley companies regularly acquire technology without ever putting any of it to use. Additionally, 2005 was a long time ago and the technology Whitman proposed in his dissertation may well be out of date today.

There is, however, good reason to believe that Spotify does in fact use Echo Nest technology today – or, that it at least shares crucial features, with respect to its attitude towards musical meaning, with the technology Whitman developed in 2005. This can be seen by closely reading the following three documents: (1) Brian Whitman's 2005 dissertation at MIT, (2) a blog post he made detailing the Echo Nest's service in 2012, and (3) the current official documentation of Spotify's API. The similarities among these three documents, which trace a timeline as long as Spotify's own, make a compelling case for the idea that Spotify's contemporary recommendation engine shares at least some features with the software originally designed by Brian Whitman in 2005. This is important, of course, because the dissertation is in the public domain and can be read in detail. Bearing in mind the important qualifications raised by Nick Seaver (2013), and being careful about the scope of our argumentation, we can ground certain claims about Spotify and automated recommendation in a close reading of the dissertation.

In 2012 (two years before the Spotify acquisition), Brian Whitman penned a blog post (Whitman 2012) outlining the Echo Nest's general approach to music information and his own opinions on the industry as a whole. This post explicitly links the technology of the Echo Nest to the research activities of both himself and Tristan Jehan at the MIT Media Lab, and most of the features he describes in the blog also appear in his doctoral dissertation. For example, in the blog post, Whitman expresses his deeply held conviction that musical similarity derives from "cultural" meaning, not simply audio signals:

> We've shown over the years that people's expectation of "similar" – either in a playlist or a list of artists or songs – trends heavily towards the cultural side, something that no computer can get at simply by analyzing a signal (Whitman 2012).

This idea – that musical meaning resides outside of the audio signal – is the central conceptual frame for Whitman's doctoral dissertation from 2005, which positions itself unambiguously against an "absolutist" theory of musical meaning deriving "from the signal alone." It is not an overstatement, in fact, to say that this is the whole idea of the dissertation. When Whitman promises to "learn" the "meaning" of music, what he is promising above all is to capture, and render legible to machines, the difficult and unruly "cultural" information that attaches to the audio signal – and then to combine the two information streams into a single classification system into which any music can be fed. The idea that musical meaning is not in the signal alone is the single most important idea animating the dissertation *and* the 2012 blog post. Thus, we have our first clear conceptual connection between the two.

In this same post, Whitman also refers to the Echo Nest's "Audio Analysis Engine," and even provides a link to Echo Nest official documentation of this product, prepared by co-founder Tristan Jehan. This document explains how the Echo Nest's machine listening works. That is, how their system deals with the audio signal itself (as distinct from the extra-signal "cultural metadata" so central to Whitman's intervention). The Audio Analysis engine detailed in 2012 bears unmistakable similarities to the one Spotify makes available today. The 2012 document, for example, takes in an audio signal and rates it in various ways. It can evaluate it in conventional musical ways, according to its *key*, *mode*, and *tempo*. These are standard music information retrieval tasks. The 2012 document also contains more idiosyncratic measures, however, such as the abstract musical categories of *valence*, *danceability*, and *speechiness*.

Crucially, all these same categories are available today in Spotify's "Get Audio Features" API endpoint.[23] Exactly as in the Echo Nest circa 2012, Spotify today evaluates tracks for their key, mode, tempo, as well as their valence, speechiness, and danceability. Moreover, in most cases the language of the contemporary API documentation echoes verbatim the language of Tristan Jehan and Whitman in 2012. Here is Whitman characterizing The Echo Nest's machine listening tool in 2012:

> We emit song attributes such as danceability, energy, key, liveness, and speechiness, which aim to represent the aboutness of the song in single floating point scalars (Whitman 2012).

Each of these idiosyncratic metrics (danceability, energy, etc.) is outlined in the contemporary Spotify API documentation, with each one still represented as a single floating-point scalar. For more commonalities, we can look at the way these fields are defined. Here, for example, is Jehan defining mode in the 2012 documentation:

> [Mode] indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived (Jehan 2012).

And here is Spotify defining mode in the contemporary API documentation:

> Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.[24]

Similarly, in 2012, Jehan defines the key output of the Echo Nest's audio analysis tool as:

> The estimated overall key of a track. The key identifies the tonic triad, the chord, major or minor, which represents the final point of rest of a piece (Jehan 2012).

Which has been somewhat refined in Spotify's 2019 documentation:

> The estimated overall key of the track. Integers map to pitches using

23 Accessed April 1, 2019 at https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

24 See "Get Audio Features" endpoint in the Spotify API, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key was detected, the value is -1.

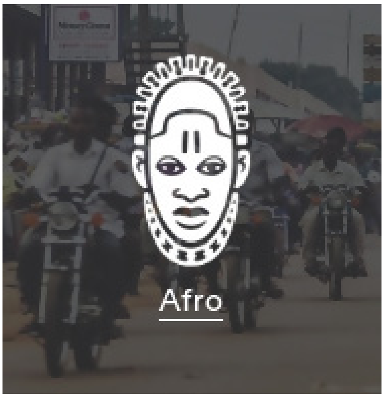
The rest of the fields exhibit the same parallelism. It seems clear that Spotify today is using the same audio feature extraction techniques that Whitman and Jehan were writing about in 2012 – which Whitman in turn explicitly connects to his own 2005 system. The grounding idea of Whitman's dissertation, moreover – that musical meaning resides not in the audio signal alone – is a prominent theme in his 2012 blog post. Reasoning from these commonalities, this article can deduce that Spotify in 2019 is still using at least some key features elaborated in 2005 by Whitman and that it is therefore likely that the "theory of musical meaning" elaborated in the one is roughly operational in the other. There is undeniably some license in this inference, and some readers may want to reject all or part of this assumption; I hope that even the most skeptical reader, however, will find the following discussion worthwhile.

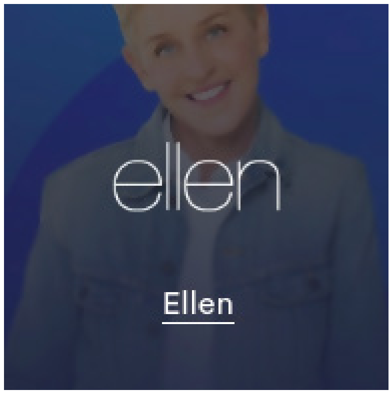
III. What does music mean to Spotify?

III (a) Spotify GUI

The theory of musical meaning I ascribe to Spotify will be principally derived from its underlying technology, which I examine mainly in the form of Whitman's dissertation (2005). Before doing that, though, it is worth taking a moment to look at Spotify's graphical user interface (GUI) to examine the notion of musical meaning implied there. Even a cursory examination of its front-end reveals some key assumptions Spotify makes about how music is meaningful to its users.

Music is grouped for Spotify users not primarily by genre or style (and certainly not by album, a concept that has grown increasingly outdated in the post-Napster world), but rather by mood, activity, and what might be termed as "musical keywords." Under the "browse" section, the user is confronted with various buttons that will lead to musical options. These are termed "hubs" in the Spotify lexicon, and they are represented by clickable square thumbnails. Hubs are distinct from the more traditional "genre" marker in that they can refer to various different kinds of musical reference. There are hubs pointing to traditional genres ("country" and "folk"), but also to activities ("party" and "chill"), as well as to politically-oriented themes ("black history is now"), sponsored content ("Spotify singles"), and, curiously, even a single hub dedicated to Ellen DeGeneres ("Ellen"). Hubs appear as thumbnail images with artwork evoking a given hub's theme (a raised fist, for example, for "black history now," a dove for "christian," and a cartoon of an African mask for the "Afro" hub). These thumbnails rework the traditional idea of an "album cover," turning it into a generic index for a given mood, more or less in the way emojis caricature human affective states. Examples of the clickable thumbnails "Ellen" and "Afro" are shown below:[25]





Although the selection of "Ellen" as a hub alongside "Afro" may seem inscrutable, the heterogeneity of the Hub themes illustrates an important feature of the kind of musical meaning the Spotify GUI seems to assume: in the curatorial phase of music streaming, music's meaning resides in its relationship to other activities or feelings. The traditional idea of *genre* is that there are certain musical properties shared among all members of a genre. The reference for a genre is, as Whitman would put it, "the signal itself." This is not true of "hubs," which are instead significant for their extra-musical references (as in the hubs "study," "sleep," "Ellen," etc.). Considered as a "hub," even the word "reggae" (apparently a *genre* word) works differently from "reggae" as a genre. Put "reggae" next to "Ellen" and you change the status of the word subtly. A "reggae" *genre* refers to the sound of the music, whereas a "reggae" *hub* is a broadly construed, fungible cultural index. Like "Ellen," it doesn't refer to a type of music so much as a musical-cultural vector.

Not coincidentally, this is exactly the kind of vector given a technical expression in Whitman's dissertation, which insists again and again that true musical meaning is informed by culture, that it is not in the audio signal alone. Of course, few people today would endorse the outmoded idea that real listening can or should take place in an idealized way, divorced entirely from extra-musical factors. Nevertheless, it is important to note that Spotify seems to have landed at the other extreme – that all music is "soundtrack." Listening on Spotify is not about attending to music but using music to evoke a desired feeling or achieve some other secondary effect. As Ellen herself puts it on a promotional web page for the "Ellen" hub, "I'm so excited to partner with Spotify on my very own music hub because music truly makes everything better. Well, music and salt."[26]

Like salt, music in the Spotify universe makes things better; presumably it

25 Screenshots taken from Spotify desktop app on May 20, 2019.
26 https://ellen.withspotify.com, accessed March 27, 2019.

also shares with salt the property of not being very good on its own. It is just one ingredient among others, one more good to consume in the effort to lead as full and happy a life as possible – something the reggae hub will help you do in a way that the reggae genre can't.

As theories of musical meaning go, this one is not crazy. The opposite extreme, where musical meaning is inherent in the idealized formal properties of a composition, is no less objectionable. It is interesting to note, however, that this is a posture Spotify arrived at mainly because it found itself having to help people discover new music; the idea of music as functional, or "relational," as Whitman sometimes puts it, is in part a byproduct of the need to make music discovery systematic and programmable. It is a music-philosophy statement arrived at because of the desperate need to accommodate a capricious market. Spotify's previous and more traditional "lean-in" posture, in which users were trusted to know what they wanted, does not rely upon any such philosophy of musical meaning. If users are finding their own music, Spotify itself is able to remain agnostic on the question of music's purpose. Users who know what they like don't want "hubs." It is only because market trends now demand a recommendation engine that Spotify has had to make choices about these questions. Its answers are visible, in part, in the user interface.

III (b) Reading Whitman, "Learning the Meaning of Music" (2005)

As Nick Seaver points out, knowing how a recommendation algorithm works is never a simple matter. Drawing specifically on his fieldwork in the music recommendation space, Seaver notes that, according to one interlocutor, there is never any single recommendation algorithm at work. Instead,

> There is not one playlisting algorithm, but five, and depending on how a user interacts with the system, her radio station is assigned to one of the five master algorithms, each of which uses a different logic to select music (Seaver 2013, 5).

When it comes to algorithms "in the wild," Seaver holds, it is never the case that they are simply a black box waiting to be opened by the right critic. The whole idea of the algorithmic black box is a red herring, a tempting fiction that tends to nourish the worst fears about algorithmic mediation. If there is a single secret code at work rather than a constantly changing and unspecifiable one, it is easy to assume the worst about it. The reality is that recommendation algorithms are far too intimately personalized, too frequently updated, and too complex for those fears to be either right or wrong in any straightforward way. This is not to say that suspicions about them are never justified, nor that the logic of a system can never be divined, but simply to remind us that we must bear in mind that our conclusions are almost always based on incomplete and possibly outdated information.

What this means in regard to Spotify, is that some types of claims are going to be more reasonable than others. We may never know how a given playlist was curated, nor, for example, what the precise proportions of "content-based" and "context-based" considerations are at work in Spotify's recommendations. But we can make empirical observations about logged recommendations, and we can think critically about the fact that every recommendation does combine the two types of signal in some way.

With these considerations in mind, one good way to approach the questions regarding the algorithm is to do a close reading of Whitman's 2005 dissertation. The technology outlined therein is distinguished above all by its ability to join two disparate subsets of music information retrieval: on the one hand, sophisticated "content-based" machine listening methods (methods that draw on the machine listening techniques alluded to above), and, on the other hand, "context-based" information culled from web crawling and other kinds of natural language processing. These two types of signal are combined into a machine learning model that, in turn, can be used to classify as-yet-unheard musical material.

Crucially, this is an approach that Whitman specifically positions against the kinds of music information retrieval techniques that derive musical meaning from the audio signal alone, which were apparently predominant in 2005. As Whitman puts it, systems that rely on the signal alone are "doomed," since they miss the essential element of human reaction. As noted above, the idea that musical meaning isn't "in the signal" is Whitman's most important theoretical commitment.

One interesting thing about Whitman's dissertation is the fact that, although it would eventually power a major corporation that many artists see as an exploitative shill for the major labels,[27] it is really an extended plea for a more nuanced treatment of musical meaning. At its heart is the kind of argument you might expect to hear from a musician: that musical meaning is hugely complex, variable, unpredictable, and contingent.[28] Whitman's language is, at times, quite personal:

> Our driving force behind this work is that fundamentally, the current approaches anger us: they don't seem right. Music is a personal force that resists 'processing,' 'packing' or 'understanding' (Whitman 2005, 91).

"Current approaches" in the above are those that take a signal-only approach (or, even worse, a context-only approach) to musical meaning. Either one, on its own, inevitably does a disservice to the true complexity of musical meaning. So far, Whitman's argument is one that probably few musicians would quarrel with. Actually, it sounds very similar to the kinds of complaints musicians frequently

27 See, e.g. (Sanchez 2018), who ranks Spotify near the bottom as one of the lowest-paying streaming services for artists, at $0.00397 per stream in 2018.

28 This is possibly because Whitman himself has performed as an avant-garde noise musician, under the stage name Blitter. According to Whitman's LinkedIn profile, Blitter's career ended in 2002. Careful not to confuse Whitman's stage name with the social network of the same name.

make of recommendation services (including Spotify): they just don't get it. But Whitman, of course, goes further than this complaint. He argues that by combining these two types of signal, one can come much closer to the true essence of musical meaning. In a basic sense, if we believe that The Echo Nest is a good system, we must agree with Whitman that he has in a non-trivial way managed to do what his title has promised: to "learn" the meaning of music. The "learn" of the title is an obvious reference to "machine learning." But what work is being done by the word "meaning?"

Newton v. Diamond and the question of musical meaning

The thesis begins by going over the well-known legal dispute between James M. Newton and the Beastie Boys over their usage of a sample from his 1978 release, *Choir*. The Beastie Boys legally licensed a few seconds of solo flute playing and looped it for their 1992 song, "Pass the Mic." The legality of the audio sample is not in dispute. Nevertheless, Newton sued for copyright infringement, arguing that the sample in question infringes upon the musical composition itself in a way not provided for by the negotiated mechanical license. Whitman uses this case to establish the central frame for his entire thesis. It proves that the true significance of music resides outside of the audio signal itself:

> When the Beastie Boys sampled his recording they took far more than the signal, even if the signal was all they took. Where can we find the rest? (Whitman 2005, 17)

After having used this case to establish the main framework for his thesis, Whitman leaves the legal questions alone. However, it is worthwhile to examine the actual facts of the case. One crucial point Whitman ignores is that the court immediately sided with the Beastie Boys. While James Newton would presumably agree with Whitman's central premise (that the Beastie Boys took more than the signal, even if it was all they took), the law does not. Strictly speaking, the only thing the case demonstrates is that James Newton *alleged* that they took more than the signal, a feeling he shares with his fellow musician Brian Whitman. Whomever we side with in the legal matter, the case does not really argue one way or another on the question of where musical meaning lives (which is Whitman's real focus in his thesis). In other words, the central frame for Whitman's "meaning," is almost off-topic.

Moreover, this case raises the issue of the "meaning" of music only in a relatively straightforward way, the same way in which almost any intellectual property dispute in music would: it points to the fact that reasonable people can sometimes disagree on what should constitute copyrightable musical property. As for the question of whether musical meaning can be convincingly derived from amalgamated reviews, Google searches, and machine learning, or whether it should be derived exclusively from the audio signal – the question to which Whitman's thesis is actually addressed – the Newton v. Diamond case bears no special relationship to it.

It is interesting to note that the case does hinge on a question of musical meaning, but one that it is different from Whitman's question. As Whitman correctly points out, the legality of the sample is not in question; the Beastie Boys obtained the rights to use the sound recording from ECM for $1000. But Newton also copyrighted the "Choir" composition, and it is this holding upon which he argues infringement has occurred. At issue, therefore, is the relationship between a sound recording and a composition, rather than, say, a listener/customer and a piece of music. The former relationship is what the legal case is about: the judges are really ruling on whether the legal instrument of a "composition" has been infringed upon by a sample deployed in a particular musical context. More specifically, what's at issue is whether or not six seconds of a sampled flute performance can constitute a vital part of the musical composition "Choir." The court upheld the verdict that, not only does the sample not constitute a vital part of the "Choir" composition, but that, even if it did, the Beastie Boys usage of it is "de minimis," that is, too insignificant to be legally actionable. As Chief Judge Mary Schroeder puts is "the dispositive question is whether the similarity goes to trivial or substantial elements."

The question is not whether the meaning can be derived from the musical stimulus but rather whether or not a small sample can infringe meaningfully upon the legal instrument known as the "composition." These are different problems. The legal case has nothing to do with the meaning of music in the broad, contextual way that Whitman will eventually construe it, that is, the sense in which music can be meaningful to a potential consumer base. Much less does it relate to the question of how that meaning can be leveraged into an effective recommendation engine. The legal case is much narrower than that, and all the argumentation connected to it remains firmly in the domain of musical form, explicitly excluding the "cultural metadata" that is so important to Whitman's work. The legal case that frames Whitman's "meaning" does tackle a problem of musical meaning, but it is not the same problem in which Whitman is interested. So, while the frame is an interesting entry to Whitman's real work, it does little to elucidate the nature of the musical "meaning" we are going to be learning about.

Whitman and Leonard B. Meyer

In spite of the critiques raised above, the case is rhetorically effective. It does seduce us into contemplating the problem of musical meaning. For Whitman, the answer is to "Learn the Meaning of Music." That is, to combine context-based (amalgamated human reactions to music) and content-based information (signal-derived) into a machine learning model that can, in turn, be used to evaluate as-yet-unheard audio signals. In sophisticated and often musically nuanced ways, ground truth data denoting the relationship of audio signal to semantic content is used to train classifiers that can determine membership of a given audio frame in a given semantic category.

At the heart of Whitman's system are machines that listen to music and, in ways informed by actual human reactions to music, determine its membership in musically useful categories. Note Whitman's usage of the idea of "meaning" in this framework:

> A model of the contextual information given a signal allows us to accurately 'understand' music (extract semantic features of link to the outside world) that hasn't even been heard yet. So what we call meaning throughout this thesis is defined as *the relationship between a signal and its interpretation*. In our work we create predictive 'machines' that analyze audio signals and extract projected community and personal reactions: these are 'meaning classifiers.'
> What we attempt to do here is computationally understand this extra-signal information and link it to the signal in such a way that it can be predicted for future audio (Whitman 2005, 19).

As noted above, this is a theory of musical meaning that Whitman posits in contrast to dominant intellectual trends in music information retrieval. The question of musical meaning is, of course, also dealt with in the disciplines of musicology and aesthetic philosophy, and Whitman situates his thesis in this intellectual tradition as well. Throughout the entire thesis, though, Whitman only cites one musicological source: Leonard Meyer's influential 1956 book, *Emotion and Meaning in Music*. This book serves as a humanistic counterexample to his own work, representing what Whitman terms the "absolutist view" of musical meaning:

> At the outset we should make it clear that our definition of meaning above is mostly referential, that is, it exists as the connection between two representations. This contrasts with the purely *absolutist* view discussed by Meyer, in which the meaning is encompassed purely within the composition or signal. Our approach considers both with an emphasis on referential types of meaning. Many musicologists study the absolutist view of musical meaning simply because there is no formal mechanism of analyzing the contextual information. What this thesis presents are ways of computationally representing both signal-derived and contextual music information and then ways of learning a model to link the two (Whitman 2005, 19).

Whitman's system combines digital signal processing techniques (content-based) with natural language processing techniques (context-based) to produce "meaning classifiers" – algorithms, trained on those two data sources, that can predict more "extra-signal information" for new, as-yet-unheard audio signals. It is a system for producing descriptions of music that incorporate both audio processing and large amounts of empirical, human-generated musical descriptions. Throughout this thesis, this system is associated with the words "meaning" and "understanding" (although Whitman sometimes places these words inside scare quotes).

Suppose that the system Whitman distills, a system that predicts "extra-signal information" about musical signals, is a good one. Whitman opposes it to Meyer's ideas, but how much distance does he really gain? In what follows I argue that the answer is "not much" – that is, that in spite of explicitly positioning himself against "absolutism" as encountered in his reading of Meyer, Whitman's approach actually aligns with Meyer's in most of the relevant ways.

Leonard Meyer serves Whitman in a similar way as the legal case discussed above. It is a framing conceit used to clarify his central intervention: that, contra both MIR and "many musicologists," meaning does not reside in the audio signal. For Whitman, Meyer exemplifies an approach to the question of musical meaning that attempts to derive it from the "signal" (from audio signal or representations in score, which, for Whitman, seem to be philosophically equivalent). "Many musicologists," Whitman tells us, take Meyer's approach, and they do so because "there is no formal mechanism of analyzing the contextual information." In other words, musicologists do not incorporate empirical human reactions into their theories of musical meaning because they lack any rigorous method for aggregating and processing them at scale. Whitman, of course, provides such a mechanism, and making this distinction is the beginning and the end of his engagement with Meyer and with the rest of the intellectual tradition for which he stands.

Whitman's system, however, in spite of its engagement with extra-signal materials ("cultural metadata") still has the same basic contour as Meyer's. Both address a scenario in which a signal is audited as the sole stimulus in a musical event. Meyer, availing himself of then-popular trends in psychology, characterizes music as a system of delayed gratification. Music sets us up to expect certain things and manipulates our innate desire to see those expectations fulfilled, in ways that stimulate complex affective responses.[29]

Whitman, as a software engineer, approaches the issue in a different way – but in spite of his protestations against "absolutism," not in a way that privileges the audio signal any less. Whitman produces a system that hears music and evaluates it, predicated on sophisticated audio- and language-processing techniques. Meyer sees musical affect as one case of a broader system of human affect, Whitman as a data science problem. Yet both authors see the process of musical meaning making as one in which the signal acts upon the listener (machine or human). Considered in this light, both authors agree on the signal as the primary source of musical significance.

Whitman's whole claim is that Meyer (and, it bears repeating, the entire discipline he stands for) fails to take contextual information into his account of musical meaning. But the truth is that Meyer does address it. Throughout his work, he is

29 The famous comparison from Meyer is that of the cigarette smoker whose emotions are piqued when he, craving a smoke, reaches into his pocket to find that he's out of cigarettes. Music, according to Meyer, triggers a similar affective response via a similar physiological mechanism (Meyer 1956, 14).

perfectly aware of the role that extra-signal information can play in the excitement of affect and construction of meaning. It is just that he regards this kind of information as outside his purview:

> We have found that the subjective data available, taken by themselves, provide no definite and unequivocal information about the musical stimulus, the affective response, or the relation between them (Meyer 1956, 12).

Elsewhere, he states this even more directly:

> Listeners and the objective data gathered from the observation of behavior and the study of the physiological responses to musical stimuli did not yield reliable information about the musical stimulus or the affective responses made to it (Meyer 1956, 22).

By "subjective data" (and, in a terminologically confusing choice, "listeners and the objective data gathered from [them]"), I take Meyer to be referring to listeners' reported affective states – the empirical responses of actual people reporting actual experiences to music. Thus, Meyer is here referring to more or less the concepts that Whitman terms "context" and "cultural metadata." For Meyer, this kind of "context" cannot tell us anything about the nature of the affective response itself, which is the essential substrate of musical meaning itself. This data is relevant to a conversation about musical meaning only in light of a *general* theory of affect, which is what Meyer hopes to explicate:

> This difficulty can be resolved only if the subjective data available... can be examined, sifted and studied in light of a general hypothesis as to the nature of affective experience and the process by which musical stimuli might arouse such experience. (Meyer 1956, 12)

First, Meyer says, you should postulate a general hypothesis about how meaning and affect arise. Then, and only then, can Whitman's "cultural metadata" figure meaningfully into a discussion of musical "meaning." Whitman is wrong that Meyer ignores human reaction because it's too difficult to integrate at scale. He just regards it as unimportant in a serious discussion of musical meaning. For Meyer, this discussion properly seeks to answer, "how does music work?" – not just "how has music worked for many people, and how best to use that information to synthesize future human reactions?"

In a part of Meyer's book that Whitman seems to have ignored altogether, this allows Meyer to imagine listening situations where context and conditioning do in fact play a large role in the construction of musical meaning. In this regard Meyer leaves much more space for extra-signal information than Whitman gives him credit for:

> Often music arouses affect through the mediation of conscious connotation or unconscious image process. A sight, a sound, or a fragrance evokes half-forgotten thoughts...These imaginings...are the stimuli to which the affective response is really made. In short, music may give rise to images and trains of thought which, because of their relation to the inner life of the particular individual, may eventually culminate in affect (Meyer 1956, 256).

He goes on to say:

> Neither the form nor the referential content of such experiences, however affective they may be, have any necessary relationship to the form and content of the musical work which presumably activated them. The real stimulus is not the progressive unfolding of the musical structure but the subjective content of the listener's mind. Yet...it seems probable that conscious or unconscious image processes play a role of great importance in the musical affective experience of many listeners (Meyer 1956, 258).

Note that Meyer here accepts the idea that "the real stimulus" *can* be something other than the signal itself. This is exactly the intuition animating the whole of Whitman's project, and it is one that he opposes, erroneously in my view, to Meyer's nominally "absolutist" paradigm. Again, it's not that Meyer ignores this fact of musical perception, but simply that he regards it as off-topic for an essay on musical meaning.

Whitman has created, essentially, a system for processing audio. It is one that is informed in creative ways by empirical human affective responses, but it is still a system for processing audio – that is, a system that grants the signal a kind of primacy. A signal goes in, a classifier does its work, and an output of some kind comes out. The nature of these outputs has certainly changed over the years, but the fundamental architecture of the system (audio in, evaluation out) is most likely the same. And insofar as that fundamental architecture remains in place, Whitman has gained no philosophical distance from Meyer, who also addresses the question of how a signal operates on a person. Meyer offers a psychological account rather than a data-driven one, but the philosophical approach to sound is pretty much the same. Whitman is correct that his approach, incorporating real human responses, is different from MIR techniques that derive from the audio signal alone. The intellectual intervention and technical innovation are legitimate (and, to judge from the success of the Echo Nest, practically effective); nevertheless, it would be wrong to locate Meyer and Whitman at opposite ends of the music-philosophical spectrum.

The real difference between the two authors, of course, is that Meyer is trying to understand how people relate to music and Whitman is building a machine that emulates how people relate to music. The machine's listening experience is qualitatively different from the human one; it is impossible for a machine to have the experience of the "subjective content of the listener's mind," to have the music call to mind a long forgotten experience which triggers a cascade of memories and affective states, or to experience listening in the company of friends. The machine "listens" in silence, in isolation, and without any subjective experience; in addressing itself to this scenario, there is a sense in which Whitman's system is infinitely more "absolutist" than Meyer's.

Although both these authors use the word "meaning," they are for the most part not on the same topic. Moreover, where their topics do overlap, they basically agree (they're equally "absolutist"). Whitman is not wrong that Meyer needs musical meaning to depend on the "signal," or, as Meyer calls it, the "stimulus." That is indeed the relationship under investigation for Meyer. Where Whitman is wrong is in claiming that this is not true of his own notion of musical meaning. For all his talk of musical meaning, on the mysterious relationship between signal and response Whitman is basically silent – and therefore gains no philosophical distance from Meyer. He simply writes about a different subject, namely, how best to simulate and synthesize that response. The essential, causal relationship between signal and response – the only question Meyer really targets, and a problem for countless other thinkers besides Meyer – is at once implicitly taken for granted and totally ignored in Whitman's project.

The Meaning of Meaning

What, then, is the "theory of musical meaning" employed by Spotify? Above I have sketched part of the answer: that music's meaning is functional rather than intrinsic, and that the mysterious ways in which music causes people to feel things – whatever they are (and Whitman definitely doesn't try to answer that) – will necessarily appear in a meaningful way *somewhere*, provided we gather enough data and treat it responsibly enough. In short, the "theory" of musical meaning is nothing more than the assumptions grounding the fields of machine learning and pattern recognition in general. As Zarsky puts it, the assumption is "that human conduct is consistent and that with sufficient data human behavior becomes predictable" (Zarsky 2016).

But is that really a "theory" at all? You might well answer "no," and you might be right. But what, then, do we make of Whitman's claim to have "learned the meaning of music?" And what do we make of Spotify's claim to be worth $10 a month? Are not both these claims grounded in the faith that Whitman and Spotify are at some level right about what musical meaning is? And is being basically *right* about musical meaning not *ipso facto* a kind of theorizing?

It is tempting to give Spotify a pass by declaring it a kind of "engineering" rather than "science." Very well, you might say, Spotify is wrong about meaning. So what? It's not a form of science, but just a collection of engineers trying to solve a problem and earn some money. But, as Pelillo et al (2015) argue, the era of machine learning has changed the way we should think about this traditional distinction:

> The scientist's occupation is seen today more modestly as a kind of problem-solving activity not dissimilar conceptually to that of the engineer, whereas on the other hand the work of the engineer is thought to produce a form of knowledge which is on a par with that produced by the scientist (Pelillo 2015).

Whitman himself dislikes the idea that man and machine stand in opposition. In answer to a reporter's question about The Echo Nest's potential to homogenize listening habits, Whitman defiantly responded:

> You call it algorithms but it's a lot more than that. We are obviously doing a ton of computer stuff but it's all based on what people are saying and choosing and that stuff. We hate this stupid man versus machine dichotomy.[30]

If the man-vs-machine dichotomy is "stupid," it should follow that the programmatically derived "meaning" is not just an engineering expedient, but a true statement about how music works for people in the real world. Pattern recognition and machine learning, in other words, are places where the line between science and engineering is blurred. The Spotify recommendation engine – whatever it really is – is in fact as much a theory of musical meaning, an attempt to characterize the process that causes people to like music, as it is a product designed to keep us logged in and spending money. In this regard it is not different in kind from Meyer, but rather in its approach to its own theoretical commitments. And, as I have shown, upon close inspection there are interesting deficiencies there.

In other words, Spotify sidesteps the question that should matter to it most (what does music mean?), even as it postulates a cryptic kind of answer (and keeps that answer a secret from its subscribers). The theory is that if we collect enough data, musical meaning, in all its manifold varieties, will be discerned by the system; as for specifying the nature of musical meaning itself, Whitman cites a single source as representative of hundreds of years of investigation into that topic, gives it a cursory reading, and then shrugs his shoulders because, after all, the real task is software design, not philosophizing.

This is a pretty dramatic intellectual liberty to take, one that Whitman is allowed because of a peculiar type of privilege he enjoys: the privilege deriving from the prestige of the discipline of machine learning, from the slippage inherent in that discipline between science and engineering, and from the financial promise of the system he created. But this privilege does not mean that the philosophical question

30 Emily White, "The Echo Nest CTO Brian Whitman on Spotify Deal, Man Vs. Machine, Why Pandora 'Freaks' Him Out (Q&A)", Billboard interview in 2014, https://www.billboard.com/biz/articles/news/digital-and-mobile/5944950/the-echo-nest-cto-brian-whitman-on-spotify-deal-man-vs, accessed May 20 2019.

shouldn't matter to the software designer; the inevitable fact is that the system's viability does ultimately depend on the way it construes musical meaning. If the "meaning" in Spotify is not the one customers value, or if it has not been "learned" in a way we are ready to accept, the whole Spotify enterprise is called into question. If meaning is as contingent as Whitman maintains, maybe another system would work just as well. Maybe *any* other system would work as well. Maybe there is no coherent way to measure how well such systems work the first place. Spotify seems to work pretty well, but so might a system of random recommendations. Given the capriciousness of musical affection Whitman mentions so often, that is a real possibility. A close look at Spotify's treatment of the problem of musical meaning reveals that it remains as obstinate a problem as it has been throughout its long history in aesthetic philosophy, a history that remains relevant even though Whitman dispenses with it in a brief passage or two. It is a problem as thorny and intractable as the financial crisis confronting the music industry in the 21st century, another problem that Spotify hasn't solved.

List of References

Aguiar, Luis, and Joel Waldfogel. 2015. "Streaming reaches flood stage: does Spotify Stiumlate or Depress Music Sales?" National Bureau of Economic Research Working Paper 21653. http://www.nber.org/papers/w21653.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. May 23, 2016. "Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks." *Pro Publica*. Accessed January 30, 2019. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

BBC News, "Newsbeat". 2018. Online Resource, https://www.bbc.com/news/newsbeat-43240886. Accessed Jan 27, 2019.

Eriksson, Maria, Rasmus Fleischer, Anna Johansson, Pelle Snickars, and Patrick Vonderau. 2019. *Spotify Teardown: Inside the Black Box of Streaming Music*. The MIT Press.

IFPI. 2018. (International Federation of the Phonographic Industry, state of the industry) https://gmr.ifpi.org/state-of-the-industry. Accessed Jan 27, 2019.

Jehan, Tristan. 2012. "Analyzer Documentation." http://docs.echonest.com.s3-website-us-east-1.amazonaws.com/_static/AnalyzeDocumentation.pdf

Kim, Tae. 2017. "David Einhorn: Value investing may be dead and Amazon and Tesla killed it", published at cnbc.com. https://www.cnbc.com/2017/10/24/david-einhorn-value-investing-may-be-dead-and-amazon-and-tesla-killed-it.html. Accessed May 21, 2019

Kot, Greg. 2009. *Ripped: How the Wired Generation Revolutionized Music*. New York: Scribner, a division of Simon / Schuster.

Lunden, Ingrid. 2014. "Spotify Acquired Music Tech Company The Echo Nest In A $100M Deal." https://techcrunch.com/2014/03/07/spotify-echo-nest-100m/

Meyer, Leonard B. 1956. *Emotion and Meaning In Music*. Chicago: University of Chicago

Press.

Pelillo, Marcello, Teresa Scantamburlo, and Viola Schiaffonati. 2015. "Pattern recognition between science and engineering: A red herring?" Philosophical Aspects of Pattern Recognition, Pattern Recognition, Letters 64:3 –10.issn: 0167-8655. doi:https://doi.org/10.1016/j.patrec.2015.06.030. http://www.sciencedirect.com/science/article/pii/S0167865515002068.

Peoples, Glenn. 2015. "How 'Playola' Is Infiltrating Streaming Services: Pay for Play Is 'Definitely Happening'." https://www.billboard.com/articles/business/6670475/playola-promotion-streaming-services. Published at billboard.com. Accessed 01-09-2019.

Powles, Julia, and Helena Nissenbaum.n.d. "The Seductive Diversion of 'Solving' Bias in Artificial Intelligence." https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53.

Putnam, Hillary. 1975. "The Meaning of 'Meaning'." *Minnesota Studies in the Philosophy of Science* 7:131–193.

Richter, Felix. 2018. "Spotify's Losses Widen as Royalty Costs Pile Up." Accessed March 20, 2019. https://www.statista.com/chart/4894/spotify-revenue-vs-costs/

Sanchez, Daniel. 2018. "What Streaming Music Services Pay (Updated for 2018)." Accessed February 5, 2019. https://www.digitalmusicnews.com/2018/01/16/streaming-music-services-pay-2018/.

Schwab, Pierre-Nicolas. 2018. "Workshop on fairness and ethics in recommendation systems." Accessed September 18, 2018. https://www.intotheminds.com/blog/en/workshop-on-fairness-and-ethics-in-recommendation-systems

Seaver, Nick. 2013. "Knowing Algorithms." Unpublished draft available on nickseaver.com. https://static1.squarespace.com/static/55eb004ee4b0518639d59d9b/t/55ece1bfe4b030b2e8302e1e/1441587647177/seaverMiT8.pdf.

–, 2018. "Captivating Algorithms: Recommender Systems as Traps." *Journal of Material Culture.*

Serafine, Mary Louise. 1980. "Against Music as Communication: Implications for Education." Journal of Aesthetic Education 14 (4): 85–96. issn:00218510, 15437809.

Silva, Marsha. n.d. "Spotify, the "Solution to Music Piracy," Is Getting Pirated by 2 Million People." Accessed February 4, 2019. https://www.digitalmusicnews.com/2018/03/26/spotify-piracy-hacked/.

Spotify 2019 home page. https://www.spotify.com/us/about-us/contact/.

Spotify 2006 home page, accessed Jan 30, 2019.. 2006. "Spotify Home Page, 2006." Accessed via the Internet Archive. Accessed April 17, 2019. https://web.archive.org/web/20061127231638/http://www.spotify.com/.

Tschmuck, Peter. 2017. *The Economics of Music*. Newcastle upon Tyne: Agenda Publishing.

Ugwu, Reggie. n.d "Inside the Playlist Factory: The Unsung Heroes of the Music Streaming Boom." Published by Buzzfeed Media. https://www.buzzfeed.com/reggieugwu/the-unsung-heroes-of-the-music-streaming-boom, accessed 5-21-2019

Whitman, Brian. "How Music Recommendation Works – and doesn't work." Blog post: https://notes.variogr.am/2012/12/11/how-music-recommendation-works-and-doesnt-work/. Accessed May 5, 2019

Yoo, Noah. 2019. "Apple Music Now Has More Paid US Subscribers than Spotify: Report," Pitchfork. Online. https://pitchfork.com/news/apple-music-now-has-more-paid-

us-subscribers-than-spotify-report/. Accessed May 20, 2019
Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.

WHAT DOES MUSIC MEAN TO SPOTIFY? AN ESSAY ON MUSICAL SIGNIFICANCE IN THE ERA OF DIGITAL CURATION.
(Summary)

This article takes it for granted that Spotify's automated recommendation engine necessarily embeds assumptions about what is musically meaningful. Given Spotify's prominence in the 21st century music industry, the contours of that theory will have definite consequences for music culture in the digital era. This article seeks to probe the latent "theory" of musical meaning underlying Spotify's recommendation technology, proceeding in three ways: first, by narrating Spotify's transition from a streaming service to primarily a "discovery" service (the so-called "curatorial turn"). Second, by making a case for why it is useful to read Spotify against the academic dissertation of a software engineer whose company it would eventually acquire (Brian Whitman's 2005 "Learning the Meaning of Music"). Third, by performing a close reading of the Spotify graphical user interface (GUI) and the Whitman dissertation, attending to the assumptions about musical meaning embedded in both. The GUI and the dissertation turn out to go well together; both seem to see musical meaning as "relational," that is, as residing in music's relationship to things outside the audio signal itself. Nevertheless there are interesting argumentative gray areas in the dissertation on the issue of musical meaning, construed as a topic in aesthetic philosophy. By examining those gray areas, this article lays the theoretical groundwork for a quantitatively derived critique of automated music curation in the future.

Article received: April 24, 2019
Article accepted: May 15, 2019
Original scientific paper