Title
Computing Taste: The Making of Algorithmic Music Recommendation

Author
Seaver, Nicholas Patrick

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Computing Taste: The Making of Algorithmic Music Recommendation

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Anthropology

by

Nicholas Patrick Seaver

Dissertation Committee:
Professor Bill Maurer, Chair
Professor Tom Boellstorff
Associate Professor Kavita Philip
Professor Geoffrey Bowker

2015

DEDICATION

To Gus

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

In writing this dissertation I have incurred many debts. First, I would like to thank my graduate cohort, who saw this project evolve in fits and starts alongside my understanding of anthropology. In particular, Chima Anyadike-Danes, with his encyclopedic disciplinary knowledge, directed my attention to many of the citations referenced here, pointing me to countless sources that have become conceptually dear to me. Beth Reddy chatted me through many tough spots, and Ben Cox never hesitated to dive into theoretical thickets.

Taylor Nelms was a clarifying light and avid reader of drafts, whose theoretical insights were invaluable. Lilly Irani pushed my thinking in fruitful directions, and Ellie Harmon commented on more drafts of this dissertation than anyone else—they both helped my writing become more adequate to the realities of making software. Jenny Carlson was a pro commiserator and advice-giver.

I am grateful to the Wenner Gren Foundation (Dissertation Fieldwork Grant 8797), the National Science Foundation's Cultural Anthropology program (Doctoral Dissertation Research Improvement Grant 1323834), and the Intel Science and Technology Center for Social Computing for their financial support. I could never have acquired that support without the help of Norma Miranda and Julio Rodriguez, and the grantwriting advice of Kris Peterson and Keith Murphy. The ISTC-Social also provided an invaluable network of colleagues and mentors, too numerous to list here.

Thank you to my interlocutors from the field, anonymized here, for enduring my questions, welcoming me into their workplaces and homes, and participating eagerly in a project whose benefit to them was always unclear.

My advisor Bill Maurer, dissertation committee members Kavita Philip, Geof Bowker, and Tom Boellstorff, along with my advancement committee members George Marcus and Melissa Mazmanian, provided helpful feedback that steered this project to where it is now. And, perhaps even more importantly, they connected me with broader networks of scholars, whose writing, workshops, and conferences have enriched my thinking tremendously. Thanks to all of my teachers. Thanks to Tarleton Gillespie and other members of the Social Media Collective at Microsoft Research Cambridge for hosting me while I finished this dissertation and worked to connect its arguments to concerns beyond anthropology.

My life changed dramatically in the course of writing up, finding me relocated across the country and setting up a new life while not yet finished with the old. Thanks to my family for helping me make the transition. Thanks to Christina Agapakis, whose work with culture(s), science, and technology has been an inspiration to me, who listened to me work out every one of the ideas contained here, and whose support made this possible. And thanks to Gus, who learned to smile just in time.

CURRICULUM VITAE

Nicholas Patrick Seaver

2007 B.A. in Literature, Yale University
2010 S.M. in Comparative Media Studies, Massachusetts Institute of Technology
2013 M.A. in Anthropology, University of California, Irvine
2015 Ph.D. in Anthropology, University of California, Irvine

FIELD OF STUDY

Anthropology of Science, Technology, and Computing

PUBLICATIONS

"The nice thing about context is that everyone has it." *Media, Culture, and Society* 37(7):1101–1109. 2015.

"Bastard Algebra." In *Data, Now Bigger & Better!*, edited by Bill Maurer and Tom Boellstorff. Chicago: Prickly Paradigm Press. 2015

"'This is Not a Copy': Mechanical Fidelity and the Re-enacting Piano." In *The Sense of Sound*, edited by Rey Chow and James A. Steintrager. Special issue of *differences: A Journal of Feminist Cultural Studies* 22(2–3):54–73. 2011.

ABSTRACT OF THE DISSERTATION

Computing Taste: The Making of Algorithmic Music Recommendation

By

Nicholas Patrick Seaver
Doctor of Philosophy in Anthropology
University of California, Irvine, 2015
Professor Bill Maurer, Chair

This dissertation reports on several years of multi-sited ethnographic fieldwork with the developers of algorithmic music recommendation systems in the US. It identifies and contributes to a nascent, transdisciplinary body of scholarship in "critical algorithm studies"—studies of algorithms' sociocultural lives by scholars outside of mathematics or computer science. It argues that critics should concern themselves not with "algorithms" narrowly defined, but with sociotechnical "algorithmic systems," of which humans are an integral part. It proposes that ethnography is a useful method for apprehending the cultural features of algorithmic systems and that these cultural features play a crucial role in the functioning of algorithms and how they change over time. Recommender systems provide a case in which to investigate these cultural concerns as they play out in the development of "preferential technics"—the intermingling of circulatory infrastructures with theories about taste. Arguing that theories of taste are embedded in algorithmic systems, the dissertation examines three areas that demonstrate this intermingling: listeners, music, and listening. The chapter on listeners describes how recommender systems have come to be used as tools for capturing users, bringing the anthropological literature on trapping to bear on the question of how imagined listeners inform the design of systems for captivating them. The chapter on music investigates

how developers imagine music to occupy a "similarity space," through which recommenders help listeners travel; theories about the nature of that space and the influence of developers on it mediate between understandings of space as a constructed or as a discovered order. The chapter on listening examines the changing techniques through which computers are taught to "hear" musical sound, arguing that the quantification of music is not simply a rationalization, but the establishment of a resonance between auditory and quantitative phenomena with unanticipated consequences. The conclusion explores the similarity between ethnographic methods and big data analytics, understood through the frame of "attention." Thinking of algorithmic systems and critical research methods as techniques for organizing attention offers new, fruitful avenues for critical algorithm studies.

INTRODUCTION:
THE ETHNOGRAPHY OF ALGORITHMIC SYSTEMS

Technology with Humanity

In August 2013, the audio company Beats Electronics announced that it was launching a music streaming service: Beats Music. Beats had become a household name with its popular headphones brand, *Beats by Dr. Dre*, and its curlicue red "b" was seemingly everywhere: in music videos, paparazzi shots of celebrities, on the heads of people on the street, and on billboards across Los Angeles, where I was living at the time. Critics had panned the headphones' technical quality—cheap components with the bass cranked up (Popper 2014)—and suggested that their success was an extraordinary branding achievement, capitalizing on the celebrity networks of the company's co-founders, Andre Young and Jimmy Iovine. Young was better known as the rapper and producer Dr. Dre, and Iovine had produced some of the most popular albums of the 1980s and 1990s. Both had since founded record labels and ascended to an industry status which led journalists to call them "moguls" or "impresarios." Iovine contested criticism of Beats' technical quality with his own musical expertise—he was "the man with magic ears," as he told a *Rolling Stone* interviewer (Fricke 2012). The *New York Times* reported:

> He dismissed those who criticize the sound quality of Beats. Competitors use fancy equipment to determine how headphones should sound, he said, whereas he and Mr. Young simply *know* how they should sound. (Martin 2011)

In promotional materials, Young argued similarly for his own expertise: “people aren't hearing all the music. With Beats, people are going to hear what the artists hear, and listen to the music the way they should: the way I do” (Burrell 2012).

Their new music streaming service launched into a similar trajectory as their headphones business had. It was also headed by a pair of music industry veterans with cultural bonafides: Ian Rogers, who had been the 1990s webmaster for rappers The Beastie Boys before running a series of digital music ventures, and Trent Reznor, the frontman of the industrial rock group Nine Inch Nails. It was a late entry into a crowded market: over the previous five years, subscription services that allowed users to stream music on demand from a large catalog had become viable businesses in the US, overcoming challenges of licensing and technical infrastructure. To take on these established services—Spotify, Rdio, and Rhapsody, among others—required deep pockets and industry connections. According to industry scuttlebutt, Beats was the only company with those resources, and it was likely to be the last.

Where other streaming services had their origins in the “tech” (i.e. computing) industry, Beats claimed its roots in the culture industry gave it a deeper, more authentic understanding of music. A job posting from the day of the announcement boasted:

> We know music - we obsess over it, and devote our lives to it. We understand music is an experience, not a utility. We realize the heart and inspiration it takes to craft music and cherish the connection between the artist and the listener. Musical taste is complex, evolving, and unique. We believe that hearing the right

> music at the right time enriches your life. It's why we're here: To deliver musical bliss, and move culture. (Houghton 2013)

Beats suggested that its music technology competitors were too much "tech," not enough "music." These competitors treated music as a "utility" and managed it with software tools that lacked cultural sensitivity, using algorithms to sort and recommend music to their listeners. According to Beats, understanding music required uniquely human sensitivity.[1] "Screw Algorithms!" read the title of a *Fast Company* blog post, "The New Music Service From Beats Uses Celebrity Curators To Show You New Music" (Titlow 2013). According to Beats, the future of music streaming was not in the technical expertise of software companies, but in the expertise cultivated by long-term members of the culture industry. The company had assembled a stable of music critics, DJs, and celebrities to compose playlists "by hand," replacing the cold robot grip of algorithms with a warm human touch.

***

"Bullshit!", tweeted Oscar, a machine learning enthusiast and head of recommendation at another tech company, "Look at their job postings!"[2]

---

[1] This kind of marketing move is by no means new: see (Seaver 2011) for an account of similar jockeying among player piano companies in the early 20th century.
[2] Throughout this dissertation, quotations from the internet have been slightly altered to protect the identity of anonymized sources. People introduced with only a first name have been given pseudonyms to protect their identity. All other quotations come from publicly available material.

By the time Beats launched, I had been conducting fieldwork for several years with people who developed algorithmic recommender systems—the kind of machinery Beats Music claimed had no hope of understanding music. My friends from the field were skeptical: if Beats was going to connect a large userbase with a large catalog of music, there was no way to do it without algorithmic help. It was absurd to suggest that these algorithms could be replaced by celebrities. What was Beats going to do, have Trent Reznor and Dr. Dre personally pick out songs for every user? It was definitionally impossible to build software without algorithms, and if Beats wanted to serve the widest possible set of listeners the most diverse selection of music, they couldn't do it "by hand." As my interlocutors pointed out, achieving that scale without algorithmic help was effectively impossible. Sure enough, when, like Oscar, I looked at the Beats Music job page, there were listings for engineers and data scientists, to be tasked with building the recommendation algorithms their press releases criticized.

When I spoke with Mike, chief scientist and long-time engineer at a well-established personalized radio service, he was frustrated. His company generated playlists algorithmically, but the algorithms relied on data that had been painstakingly produced by human experts. Depending on who was talking and when, his company was described as cluelessly technical—reliant on algorithms that couldn't understand why a person had liked or disliked a song—or hopelessly human—unable to recommend new music until it had cleared the experts' backlog. Earlier in the company's life, the scientific appeal of the recommender system had been a marketing boon, but they soon found themselves caught up in this ambivalence about human and machine capacities.

Brian Whitman, co-founder of The Echo Nest, a "music intelligence" company that had come to represent the power of algorithms to parse musical data, recognized this tension as "the postmodern insanity of a *computer understanding how should you feel about music*" (Whitman 2013; emphasis in original). Journalists on the music technology beat never seemed to tire of the "humans vs. algorithms" frame, but my interlocutors, the humans cast onto the "non-human" side of the comparison, had grown weary of it. As Whitman wrote: "This is somewhat unfair and belies the complexity of the problem. Yes, we use computer programs to help manage the mountains of music data, but so does everyone, and the way we get and use that data is just as human as anything else out there" (Whitman 2013). When Spotify, which had acquired The Echo Nest, released a new recommendation feature two years later, product manager Matthew Ogle argued: "For describing the way we do things at Spotify, 'human vs algorithm' doesn't even make sense anymore" (Dredge 2015). Although the recommendations obviously depended on algorithmic processing, Ogle maintained that the new feature was "humans all the way down [...] Our algorithms stand on the shoulders of (human) giants" (Dredge 2015). While catching up with, Richard, a long-time interlocutor, in late 2015, he told me that the popularly imagined antagonism between humans and computers didn't make sense to him: "The computers help the people."

***

I had entered the field interested in precisely this tension: as music streaming services grew, algorithmic recommender systems grew alongside them, taking on the apparently paradoxical work of accounting for taste. My interlocutors found themselves caught up in this discourse about culture and technology, humans and algorithms, as they tried to explain their work — to themselves, to the public, and to me, as I conducted fieldwork with them between 2011 and 2015. I wanted to know: How did the people who built these systems reconcile the dominant perceptions of technology as essentially rational and objective and music as essentially expressive and subjective? What was happening to the relationship between taste and technology with the advent of these systems that seemed to ignore common sense ideas about their incompatibility? If practitioners in the field were fighting against "human vs machine" narratives and emphasizing the influence of humans over the propensities of algorithms, then what kinds of systems were they building, and according to what theories about music, listeners, listening, and computation?

Although these claims about the importance of humans—from companies like Beats to companies like The Echo Nest—came at a moment when public sentiment seemed to be turning against algorithms, they should not be understood as simply PR spin. The point is simple, but often crushed beneath oppositional understandings of culture and technology: no matter how technical they seem, software systems always, though in varying ways, require humans to function. The question was not whether a given service was suitably human or algorithmic, but rather how humans and algorithms interacted and what capacities were ascribed to both of them. We can hear about the necessary interrelation of culture and technology, humans and algorithms, music and machines, from professionals working in the field; we can expect it from the literature in anthropology and science and technology studies; and we can observe it directly, as I did during my fieldwork. As we will see, this interrelation is not merely a matter of determining the domains proper to humans and algorithms and then allocating the work accordingly—the "human" and the "algorithm," and their capacities, are produced through their interrelation. The persistence of the idea that these areas are somehow opposed—both in popular discourse and in certain strains of academic criticism of algorithms—is thus something of a puzzle. It was part of the cultural milieu in which my interlocutors had to explain and perform their work, and it was likewise a commonsense understanding that I have had to reckon with in my own work.

In this dissertation, I pursue a more expansive reading of what constitutes an "algorithm," agreeing with my interlocutors that the humans involved in these systems play critical roles in shaping how they work. Contrary to the common tendency to imagine that the trouble with algorithms is that they are too simple and too inhuman, I examine how algorithmic systems are constituted not only by narrowly defined "algorithms," but by people, cultural norms and institutions, developers' proclivities, and, crucially, by how these forces interact with and shape the ways that developers *pay attention* to their objects. In this introductory chapter, I outline the contours of this debate in the emerging research area of critical algorithm studies and describe my construction of a field site in which to investigate these questions. In the next chapter, I turn to the production of algorithmic recommender systems specifically, describing their origins in persistent ideas about the limits of human attention. I then outline my understanding of them as a form of "preferential technics," intermingling technical infrastructures with theories about taste. In the following chapters, I examine this intermingling in three areas: understandings of listeners, music, and listening. Taken together, these chapters explore how systems come to work as they do in the interstices of taste and technology. Rather than simply rejecting the facile opposition of technology and culture, replacing it with the vague claim that everything is sociotechnical, these chapters seek to describe how particular understandings held by humans come to matter for the functioning of algorithmic systems. These understandings address the nature of taste, the variability of music, and the salience of musical sound, and they all work to chart middle paths between the supposed opposition of computers and humans.

***

Before it was "Beats Music," it was known by the codename "Daisy,"[3] and its executives seemed more open to technical solutions to the problem of recommending music. I met Luther, one of the lead engineers on the Daisy project in early 2013, at a "music technology summit" in San Francisco—a semiannual meeting for entrepreneurs whose businesses and business plans involved music and computers. He was secretive about his work but happy to tell me about his PhD in physics, in the controversial subfield of string theory. Such academic backgrounds were not uncommon among the emerging class of "data scientists" in the Bay Area, employed by startups in many domains to help glean information from large data sets. Company founders often talked about how many "PhDs" they had, referring to their employees by alma mater and capitalizing on the academic cachet of institutions like Stanford, MIT, and Carnegie Mellon to sell potential investors on their technical credibility. When I asked Luther what string theory had to

[3] As in "Daisy Bell," the 1892 song first sung by an IBM computer in 1961 and then by the malfunctioning artificial intelligence HAL 9000 at the end of Stanley Kubrick's *2001: A Space Odyssey*. This was, according to Daisy executives, a coincidence (Van Buskirk 2013).

do with music recommendation, he told me, opaquely: “Well, they both involve dimensionality reduction.” This evasive answer, hinting at some generalizability I didn’t quite understand, was indicative of the layers of secrecy and technical obscurity my fieldwork put me up against.

By the time Beats Music was announced later that year, industry scuttlebutt indicated that this engineering team had been fired, and it was their jobs that were posted on the Beats website. This was a “very music industry move,” one of my interlocutors told me: where a tech company would gradually iterate from an early prototype, an executive in the music industry would rather fire everyone, wipe the slate clean, and start again. While early publicity around Daisy had trafficked in the potential of obscure technical achievement—the kind of thing an engineer might tease with oblique references to theoretical physics—now Beats Music would be committed to a more human touch, but looking for engineers on the side.

Behind schedule, Beats Music launched in January 2014. The app’s home screen, titled “Just For You,” displayed a set of playlists, composed by Beats’ experts and recommended by their algorithms, drawing on a list of favorite genres and artists requested from the user when they first logged in. The company released a 90-second promotional video, in which Reznor read a manifesto over a scratchily animated series of red-on-black images—silhouettes kissing, turntables spinning, a sailboat tossed on stormy water that transforms into a field of 1s and 0s. The manifesto was a pastiche of clichés about humanness and musicality:

What if you could always have the perfect music for each moment, effortlessly? Drives would be shorter. Kisses, deeper. Inspiration would flow, memories would flood. You'd fall in love every night. And life would be infused with magic.

If you want to conjure that power for people, you'd first have to respect it, be in awe of it, and realize music is much more than just digital files. It breathes and bleeds and feels.

And to do that you'll need more inside your skull than a circuit board. Because code can't hear the Bowie in a band's influences. It doesn't know why the Stones segue perfectly into Aretha Franklin.

And if you're one perfect track away from getting some satisfaction, you'd want more than software to deliver it. You'd want brains and souls. You'd want people driven by a passion for music, who know the only thing as important as the song you're hearing now is the song that comes next.

So that's what we've done. We've created an elegant, fun solution that integrates the best technology with friendly, trustworthy humanity—that understands music is emotion, and joy, culture… and life.

This is a completely new way to experience music and the next step in the evolution that's taken us from 45s to CDs to streaming. But the most important thing about it is that you'll be blown away by what happens when you hit play.

The right music. Like magic.

Though this new publicity push maintained the importance of human qualities—breath, blood, feeling—it offered a more ambivalent message about how those qualities related to technology. Beats was not solely human—as my interlocutors pointed out, it couldn't be—but rather an integration of the "best" human and technological traits. An online ad campaign lifted the imagery and sentiment from Reznor's manifesto: a heart plugged into a circuit board coursing with blood, and written across the center: "Technology with humanity."

Critical Algorithm Studies

These debates in the world of music recommendation are part of a broader popular discourse about algorithms. Over the past decade in the US and Europe, algorithms have slid into public consciousness, becoming objects of concern for people outside of computer science and software engineering. These increasingly conspicuous algorithms sort search results for Google, they filter posts on Facebook, they discern trends on Twitter, and they recommend music to listeners on Pandora. Their less conspicuous relatives help determine eligibility for loans, direct cop cars toward possible crime hotspots in predictive policing programs, and flag suspicious activity for government surveillance. As more and more of human life is conducted alongside and through computer systems, contoured by algorithmic selection both online and off, people who once had little interest in the workings of computers now have a growing interest in their effects. This attention has manifested in popular books (e.g. Steiner 2012; Carr

2014; Dormehl 2014), newspaper columns (e.g. Wortham 2012; Singer 2014; Wieseltier 2015), blog posts (e.g. Hill 2011; Friedersdorf 2014; Elkus 2015), and a series of high-profile scandals about the functioning of algorithms at large tech companies (e.g. Grimmelman 2014; Tufekci 2014).

In the academy, this public interest has been matched by the growth of a body of work I have taken to calling "critical algorithm studies." This research approaches algorithms from a variety of angles defined not by the concerns of computer science, but by those of the humanities and social sciences. During the research and writing of this dissertation, many conferences and panels dedicated to the social life of algorithms appeared ("The Politics of Algorithms," 4S Copenhagen, 2012; "Governing Algorithms," NYU, 2013; "The Contours of Algorithmic Life," UC Davis, 2014; "Algorithms and Accountability," NYU, 2015). Critical algorithm studies is transdisciplinary, spanning cultural studies (Striphas 2015; Andrejevic 2013), critical theory (Galloway 2006; Parisi 2013), law (Pasquale 2015), communication (Cheney-Lippold 2011; Napoli 2014; Gillespie 2012, 2014; McKelvey 2014), media studies (Beer 2009, 2013; Nakamura 2009; Uricchio 2011; Mager 2012; Bucher 2012; Mahnke and Uprichard 2014), journalism (Anderson 2012; Diakopoulos 2013), history (Ensmenger 2012), geography (Lyon 2003; Graham 2005; Amoore 2011), sociology (Snider 2014; MacKenzie 2014) and anthropology (Helmreich 1998; Asher 2011; Seaver 2012; Kockelman 2013). If we include "big data," the massive, typically unstructured databases that require algorithmic tools to be made tractable, then this glut of literature increases tenfold.

This critical interest in algorithms is animated by two primary concerns: The first is what happens when the rigid, quantitative logic of computation tangles with the fuzzy, qualitative logics of human life. Like algorithms, this concern is not new: similar questions have gone by the names "rationalization" or even "modernization" before. Algorithms require or produce the formalization of informal qualities: they take uncertain, personal, situational things like your taste in music or your political sensibility and render them as stark quantities. This encounter between culture and technology seems to invite grandiose claims about intrinsic "cultural logics" of broad and heterogeneous technical formations, even as a significant and growing body of scholarship in STS and allied fields undercuts the assumed opposition between technology and culture, emphasizes the emergent properties of technical systems in use, and points to the importance of historical and contextual specificity in understanding sociotechnical arrangements.

The second concern is that the most influential algorithms perpetrating this rationalization are typically obscure. They are hidden behind corporate secrecy, they require technical training to understand in detail, and they can be so complex that they pose interpretive challenges even to the people who make them. As Frank Pasquale exhaustively documents in *The Black Box Society* (2015), algorithms are often (though not necessarily) black boxes, their interiors invisible to and unalterable by those they impact, and their obscurity maintains extensive power imbalances. Companies are loath to open up these black boxes, worried that revealing their "secret sauce" will enable bad-faith actors to game the system and aid corporate competitors (Granka 2010).

The expansion of interest in algorithms has been accompanied by an expansion in the definition of the term "algorithm." Though the use of "algorithm" by critical academics signals a renewed concern with technical detail—something more particular than broad arguments about "the digital" or "new media"—many scholars are guilty of terminological fuzziness. It is not uncommon to read critical papers about "algorithms" that focus on features that would more properly be considered part of the user interface, for example. In this fuzziness, scholars participate in a common popular use of the term "algorithm," which Tarleton Gillespie has identified as a kind of synecdoche—the use of a part to stand for the whole—signifying broader concerns about computation and cultural life (Gillespie 2014). People may say "algorithm," but they mean "all that stuff Facebook is up to with the computers, including the algorithms proper, but also many other things."

One response to this situation is to call for terminological precision, narrowing in on what algorithms actually are and leaving behind the somewhat embarrassing definitional mistakes of the recent past. Computer science provides a number of technical distinctions for talking properly about algorithms, and we might use them: they are to be distinguished from the data structures on which they operate; their theoretical expression is independent of their implementation in any particular programming language; and formally defined algorithms are distinct from ad hoc heuristics. Algorithms proper, as my interlocutors in the field and (occasionally) in critical algorithm studies remind me, are ultimately math—they are just a sequence of steps—they are just a recipe—they are just "logic + control" (Kowalski 1979). These distinctions can be useful for building computational systems, and they are central elements of the emic understanding of computation among software engineers.

Critics often assume that algorithms are simple, straightforward things, following a textbook definition, like this one from a popular textbook: "an algorithm is any well-defined computational procedure that takes some value, or set of values, as input and produces some value, or set of values, as output. An algorithm is thus a sequence of computational steps that transform the input into the output" (Cormen et al. 2009, 5). Indeed, their simplicity and straightforwardness—their need to first transform any problem into "a well-specified *computational problem*" (Cormen et al. 2009, 5)—are why they are imagined to have problems grasping the complexities of social life. These understandings of algorithms have shaped critical approaches to producing knowledge about them: because algorithms are secret, knowing them is a matter of uncovering secrets. I call this the "revelation model" of knowledge about algorithms, which undergirds a set of methods: calls for transparency (Introna and Nissenbaum 2000), reliance on patents and press releases, audits of algorithmic functioning (Sandvig et al. 2014), or attempts to reverse engineer them (Diakopoulos 2013; Gehl 2014). While the scholars behind these approaches typically agree that humans have effects on algorithmic functioning, they tend to focus on algorithms in a narrow, technical sense, as matters of inputs, procedures, and outputs that can be analyzed for bias. Algorithms may manifest the biases of their creators in the sense that technologies have politics built into them (Winner 1986), but what makes them distinctive, in this view, is that they rationalize and extend these biases autonomously.

However, critical algorithm studies has several good reasons to reject the Computer Science 101 definition of algorithm as a frame and to seek out new understandings of what algorithms *are*. As I found when I went looking for them in the field, textbook algorithms only exist in textbooks. Out in the world, algorithms do not operate independently from data structures but are carefully tuned to the particularities of the data they work on and with; they are always implemented in some programming language, introducing the possibility of errors and unforeseen entanglements with computing architecture; and although a given algorithm may be well defined, the choice among algorithms was likely made according to ad hoc heuristics. While my interlocutors would defend the merits of the textbook definition of "algorithm," they would *not* agree that the features outlined in *Introduction to Algorithms* govern the "algorithms" they worked on. They were, as a rule, quite aware that their judgment played a role in shaping these systems—after all, exercising this judgment was their job. Algorithms, regardless of how they are defined, are emplaced in social and cultural worlds. In other words, the term "algorithm" is what Otto Neurath called a *ballung*, or congestion, "a linguistic manifestation of the ineliminable social, everyday, non-theoretical vague elements of language and of the intrinsic complexity of the world" (Cat 2014). Even among theoretical computer scientists, it has proven ironically difficult to find a precise and comprehensive definition of "algorithm," as "the notion is expanding" (Gurevich 2011).

The black box metaphor and narrow technical definition should not fool us into thinking that algorithms are simple, discrete, or straightforward things. Although their input-output functionality and obscured interior is the archetype for the "black box," these boxes are not sealed. The conspicuous algorithms that catch the attention of critical algorithm studies are typically implemented on networked computers, subject to constant manipulation by people.[4] "The more automatic and the blacker the black box is, the more it has to be *accompanied* by people" (Latour 1987, 137). Especially when it comes to the kinds of algorithms that capture the most attention these days—the ones that work on the Facebook or Google servers, for example—we cannot imagine that an "algorithm" is a simple, straightforward black box. At any minute, the algorithm you are interacting with has adjusted its function, personalized to you, it is one of many being tested simultaneously across the userbase, and it is different from the algorithm that was implemented last week and the one that will be implemented next week. You can't log into the same Facebook twice.

These algorithms are not standalone little boxes, but densely contextual ones, worked on in sociocultural worlds with hundreds of hands reaching into them, tweaking and tuning, swapping out parts and experimenting with new arrangements. They are full of people. If we are worried about algorithmic logics, we will not find them in the code, but in the motion of the hands and the broader contexts in which they work, choosing among algorithms and data representations, mediating between ideas and implementations. It is not the algorithm, narrowly defined, that has sociocultural effects, but all of this stuff—what I will call *algorithmic systems*—intricate, dynamic

[4] This excludes a class of algorithms that are implemented and then "left alone," such as fly-by-wire autopilot systems, or algorithms that are very rarely changed, like those used by governments for calculating expenditures based on census data. While there are convincing arguments to be made that those too have sociocultural lives, the present case requires even less work to argue: humans are "in the loop" constantly and

arrangements of people and code. Outside of textbooks, "algorithms" are almost always "algorithmic systems."

This shift in focus has consequences for understanding the relationship between technology and culture, and in particular the idea that algorithms are not culture but have negative effects on it. Rather than starting our critiques from the premise that we already know what algorithms do in relation to culture—they reduce, quantize, sort, and control—I take the operation of algorithms as a research topic. How do practices within algorithmic systems define and produce distinctions and relations between technology and culture? With this as our guiding question, the relationship between "outsider" humanistic or social scientific critics and "insider" engineers changes. When all there is to know about an algorithm is its function and effect, then expertise is neatly split: engineers know about functions, social scientists about consequences. But, if we think of algorithms as algorithmic systems, a field opens up for engineers and social scientists to critically engage with algorithmic systems together.

Algorithms, as Laura Devendorf and Elizabeth Goodman (2014) have argued, invoking Annemarie Mol (2002), are *multiple*—they are hard to pin down in practice, and they are enacted differently by different groups of people, from database engineers to product managers to end users. Different methods of apprehending algorithms enact them differently, and these variations bring certain features of the algorithm in and out of focus. My understanding of algorithms as algorithmic systems is thus tied to the

---

influentially. Thanks to Donald MacKenzie and Baki Cakici for bringing these examples to my attention.

method I used to apprehend them: ethnography. Ethnographic methods, I argue here, help us gain purchase on topics that concern scholars of critical algorithm studies: the interrelation of people and computers, the production of formal representations, and the sensemaking logics that lie behind algorithmic operations, locating these within sociocultural contexts that are variable and changing.

When we realize that we are not talking about algorithms in the technical sense, but rather algorithmic systems of which code sensu stricto is only a part, their defining features seem to reverse. If algorithms are formal, rigid, and consistent, algorithmic systems are often in flux, revisable, and subject to informal negotiation. If algorithms seem strictly technical, algorithmic systems spread across institutional settings, incorporate explicitly cultural theories about how the world works, and grow and change in social contexts. Where algorithms close into stable, discrete, obscure black boxes, algorithmic systems open out into the world, incorporating bits of code from elsewhere, relying on an archipelagic geography of physical computers, software services, and data sources, which flow and change constantly, challenging developers' patience, but also providing the dynamism and potential for constant improvement that tantalizes contemporary programmers. When our object of interest is the algorithmic system, "cultural" details *are* technical details—the tendencies of an engineering team are as significant as the tendencies of a sorting algorithm.

This means that if we want to understand how algorithmic systems work in relation to people, it is not enough to read patents and press releases or to experiment at the interface: we need to examine the practices at work within them. These practices blend "technical" and "cultural" concerns, and they are located in specific sociocultural contexts. This means that, in spite of what a narrow understanding of "algorithm" might lead us to believe, algorithmic systems are choice objects for ethnographic study, which examines the everyday situations in which they develop. You can visit an algorithmic system, talk with the people who work there, look at lines of code, and record the imponderabilia of actual life within them: "phenomena which cannot possibly be recorded by questioning or computing documents, but have to be observed in their full actuality" (Malinowski 1922, 18).

The Ethnography of Algorithmic Music Recommender Systems

Music recommender systems provided a useful case to examine the concerns of critical algorithm studies for a number of reasons. The first is that music itself embodies the paradoxical relationship between culture and technology that concerns about algorithms emerge from. As Georgina Born reminds us, music "destabilizes some of our most cherished dualisms concerning the separation not only of subject from object, but present from past, individual from collectivity, the authentic from the artificial, and production from reception" (Born 2005, 8). I would add "between culture and technology" to this list. Music, a paradigmatic example of "culture" by any definition, has always been technical as well, although its technical mediations have moved among a number of "assemblages," as Born terms them: from collective playing of early instruments, to written notation, to the concert hall and virtuoso performance, to audio recording, to DJing, to large catalog on-demand streaming services, and so on. Music is already a hybrid of the technical and the cultural, and this has long been a matter of some unease, especially in the Western tradition, with its trademark anxieties about mixture and impurity.

Second, people tend to not take music seriously. From dismissive assessments of pop stars to dominant ideologies about the inconsequence of entertainment media, music is not considered a "serious" domain to work in. The evolutionary psychologist Steven Pinker, for example, calls music "auditory cheesecake" (1997, 534), incidental to the driving forces of human existence. Pierre Bourdieu, in *Distinction*, seizes on music for his study of taste, describing it as "the 'pure' art par excellence"—perfect for the expression of arbitrary social distinction—because "it says nothing and has nothing to say" (Bourdieu 1984, 19). My interlocutors who felt there was more to music than this often experienced this dismissive attitude as well. Academic researchers were familiar with the fact that many of the techniques used in music recommendation, such as machine listening or matrix factorization, had derived from other applications that were considered more prestigious. Computer vision and abstractly theoretical machine learning research held more clout than their work that was often defined by its applications to music recommendation.

Although these dismissive assessments have been convincingly dismantled by scholars of popular culture and music more specifically (e.g. DeNora 2000), this understanding of music remained a feature of the cultural world in which my interlocutors operated. The "unserious" position of music recommendation relative to other applications of algorithms had useful consequences for my research. When it came to recommending music, the developers of algorithms felt a flexibility that is elusive in domains like credit scoring or medical analytics. Lives were not hanging in the balance, and as a result, music provided a space for social experimentation. Developers were also more willing to talk about their work in detail and to speculate on possible futures, not hemmed in by the worry that they might inadvertently endorse racial profiling or government surveillance. (Though, as I found, these taboo classificatory practices lurked around the edges of the work anyway, and developers were quite conscious of avoiding them.) That music offers a space for social experimentation is partly Jacques Attali's argument in *Noise* (1985): music has a "prophetic" function, anticipating broader social changes, because it offers a space in which alternative arrangements of society and culture can be temporarily taken on—attempts to break free of the historically dominant mediating assemblages Born (2005) describes.

The work of making algorithmic music recommendation is spread across a range of sites. This is not only because many companies compete or many academic researchers pursue related projects, but also because recommenders themselves draw together elements produced, collected, designed, and run in diverse locations. A music streaming service might acquire files from a digital media distributor like 7digital. It might then process metadata provided by record labels or a company like Gracenote to organize the music and display information about it to the user. Then, it might use the recommender API (application programming interface, a way for software systems to communicate with each other) of a company like The Echo Nest to generate playlists based on listener data they collect from their own platform. It can then reconcile this with patterns in online chatter, listening behavior, and the audio waveforms themselves. All the components of this system are variously available for alteration, inspection, and contestation, depending on where you are located.

To investigate even one algorithmic system thus requires the ethnographer to move, following flows of data, code, and business ventures across multiple sites. This poses an additional challenge in that many of the sites in question are protective of their intellectual property, concerned about their reputations, and unfamiliar with ethnographic work, understanding the ethnographer as someone analogous to a muckracking reporter. These challenges are not unique to the study of algorithmic systems—anthropologists have dealt with secrecy (Jones 2014) before—but they are pronounced and formalized at corporate sites through the extensive use of non-disclosure agreements. In tracing the outlines of an algorithmic system, then, one is likely to encounter dead ends, refusals (Ortner 1995), and absences. While these patterns of knowing and not knowing, of access and refusal, can be instructive in themselves (Jensen 2010)—the number of users a service has, for example, is jealously guarded, as are the terms of relationships between streaming services and record labels—many of the practices of interest to critical algorithm studies are not outright hidden, but rather obscured (see Mahmud 2012 on discretion).

This means that it is possible to see quite a bit more than is commonly assumed, but it requires a trade-off: an outsider is unlikely to be able to trace out the bulk of a single company's algorithmic system, but in roaming across the broader field of research and development, she may be able to piece together elements from disparate sites to get a broader picture of how algorithmic systems are put together. The resulting knowledge is not an exposé of a particular company's configuration at one historical moment, but rather a more generalizable appraisal of "home truths" (Geertz 1984) that obtain across multiple sites within an industry. These understandings are more durable and more common than particular technical configurations, and thus arguably more useful to outside critics than knowledge about what is inside one rapidly shifting black box at one moment in time.

To get at the patchy obscurity of algorithmic systems, I resorted to a technique used by Hugh Gusterson in his study of nuclear scientists, which he termed "polymorphous engagement":

> Polymorphous engagement means interacting with informants across a number of dispersed sites, not just in local communities, and sometimes in virtual form; and it means collecting data eclectically from a disparate array of sources in many different ways. (Gusterson 1997, 116)

The polymorphous ethnographer is a scavenger, retaining the "the pragmatic amateurism that has characterized anthropological research" (Gusterson 1997, 116; see also Riles 2001 on amateurism) and gleaning information wherever possible—in off-the-record chats with engineers about industry scuttlebutt, by reckoning press releases against the social media updates of his interlocutors, through interviews where people spill the beans and those where they give their well-rehearsed spiels, in conference hallways, and in classrooms. "Ethnography," as Ulf Hannerz writes, "is an art of the possible" (Hannerz 2003, 213).

For Gusterson, the polymorphous ethnographer is something like the "outlaw ethnographer" (Pierce 1995), tasked with freeing information from behind the walls that protect it (though, as Gusterson notes, an antagonistic relationship with one's informants—and in the combative style, "informant" seems the more accurate term — tends to serve the ethnography poorly). This imagining of the field resonates with what Emily Martin (1998) describes as the "citadel" model of the anthropology of science, in which the role of the ethnographer is to pierce the walls that scientists have set up to separate themselves from society. This serves to leak out information that is hidden— about the day-to-day practices elided in formal accounts, for example—as well as to create a new image of science in which it is more continuous with the sociocultural terrain in which its citadel has been established. But the "agonism" intrinsic to this approach tends to overemphasize the coherence inside the citadel (Martin 1998, 28), as though insiders and outsiders had descended from different branches of a segmentary lineage system, temporarily unified only for the purposes of the fight (Evans-Pritchard 1940). Efforts at breaking down the walls of the citadel are like efforts at opening black boxes. For the present case, it is as though the blackness of the black box and the impenetrability of the corporation walls are partly constituted by the agonistic approach itself—alternative images of the knowledge project at hand provide new kinds of access.

Martin describes how, especially within feminist STS, the masculinist work of storming the citadel[5] has been replaced by the tracing of rhizomatic connections, the "discontinuous, fractured and nonlinear relationships between science and the rest of culture" (Martin 1998, 31). Recognizing these relationships provides ways to understand what is going on in obscure systems, and can prove *more* useful for purposes of critique: rather than discovering precisely what Facebook is doing right now, we learn the connections that will continue to shape what Facebook does in the future, and which also shape what other, non-Facebook companies do.

But the most promising metaphor, which Martin picks up from Donna Haraway, is that of the "string figure"—the figurative game played with loops of string, documented around the world, but relegated to a curiosity (see, e.g., Jayne 1906).

> [String] figures can be passed back and forth on the hands of several players, who add new moves in the building of complex patterns. Cat's cradle invites a sense of collective work, of one person not being able to make all the patterns alone.... It is not always possible to repeat interesting patterns, and figuring out what happened to result in intriguing patterns is an embodied analytical skill. (Haraway, quoted in Martin 1998, 36)

The making of string figures is a technique (in the sense of Mauss 1973), it is figurative, although the figures often require training to interpret, and these figure-techniques are passed between people in tension, requiring active work to transmit and always changing in the process. As Martin concludes of the relationships across what is discretized as "science" and "culture":

---

[5] The gentler but no less gendered counterpart to this knowledge metaphor is the unveiling of hidden knowledge, the stripping of the black box to reveal the secrets inside (see, e.g. Harding 1986).

> Some links are invisible and disappear from time to time below the surface of what we can know into dreams, memory, or the account books of multinational corporations. Like string figures, culture is nonlinear, alternately complex and simple, convoluted and contradictory. As often as not, its processes celebrate mystery and opacity. (Martin 1998, 40)

The polymorphous ethnographer is a participant-observer among figure-techniques, tracing their lines, finding the places where they are passed around, transmitted and changed, and documenting how they are understood by those who make them. The resulting field is multi-sited in the conventional sense (Marcus 1995; Hannerz 2003), but it is also a heterogenous network of people and techniques "defined not only by social networks but by material flows and other modes of connection" (Burrell 2009, 191).

I thus constructed the field by following the making of music recommendation through a network of academic and industry researchers, primarily based in the United States, where most of the large commercial music recommenders in the world today are built. Between 2011 and 2015, I visited academic labs and corporate offices in San Diego, Los Angeles, San Francisco, Boston, and New York, for terms ranging from single lab meetings to a three month internship at a music recommendation company. I attended hackathons, meetups, and social gatherings in and out of work. I attended 9 international conferences with my interlocutors on topics including recommender systems research, music informatics, and applied mathematics. I took courses in recommender system design online and at my home institution, UC Irvine. While in the field and back at home, I kept up with my interlocutors through Twitter, mailing lists, and blogs, which provided a critical site for keeping track of work that was widely dispersed and staying in touch with individual interlocutors whose geographical movements did not neatly track with mine. I built an archive of news coverage, company white papers, and academic research on music recommendation. And, over the course of the research, I conducted 89 semi-structured interviews with CEOs, aspiring founders, product managers, software engineers, data analysts, data curators, office managers, consultants, interns, hackers, scientists, graduate student researchers, and professors. Although the protections afforded by my IRB prevent me from disclosing which companies these people worked for, they represent nearly every major music streaming service operating in the US at the end of 2015.

In the process of constructing the field, I began to recognize the contours of an "invisible college" (Crane 1972) of researchers who had begun their careers in academia and then moved out into industry, often drawing out their students and co-authors after them. These researchers passed the figure-techniques of algorithmic recommendation among themselves. Many of them had been affiliated with ISMIR, the International Symposium on Music Information Retrieval, which since 2000 has been the primary conference for computer scientists interested in computational techniques for understanding and arranging music. Many of these techniques posit music recommendation as their "real-world" application, and as I attended three annual ISMIR meetings in Miami, Porto, and Taipei, I saw a shift toward research engaged with industrial music recommendation applications in the accepted papers and how they were discussed. Researchers from ISMIR founded their own companies, were hired into existing music streaming companies, and they shared data and techniques with their colleagues working in industry. So although the techniques I saw presented at ISMIR were not necessarily those in use inside of companies (often, due to differences of scale between academic and industry datasets, academically developed techniques were simply unusable in commercial applications), they existed in the same milieu and the subset of ISMIR researchers with whom I spent most of my time shared sensibilities with their industry colleagues whose work was kept under wraps.

Pursuing this fieldwork presented a few methodological challenges:

Interview-centricity

In his essay on multi-sited ethnography, Ulf Hannerz notes that multi-sited projects are often more dependent on interviews that "traditional" ethnographies, attributing this in part to the shortened time frame allowed for each site (Hannerz 2003, 211). I found this to be the case as well, not only because my movement across the field made establishing the patterns of participant observation challenging, but because interviews were a kind of social interaction my interlocutors understood well.

My interviews were fit into well-defined periods of the day: the coffee date, the office lunch, the default hour reserved through a company's calendar system. The sociotechnical structure of meetings became a stumbling point for me, as I would send a nervously worked over email to someone requesting an interview and they would reply, "Could you send me a meeting invite by any chance?", referring to the scheduling technology that was necessary for them to fit my request into their calendared day. Interviews were thus not an occasion for me to remove people from their daily work so much as they were occasions for my interlocutors to format my fieldwork efforts into a shape that fit the normal course of the workday, as they would with any of the countless other one-on-one meetings that dotted their calendars.

As Jenny Hockey has argued (2002), interviews are not necessarily opposed to participant observation, pulling people out of their ordinary lives into a constructed, interviewer-controlled environment. In my field sites, one-on-one meetings and interviews were a deeply familiar social form. My interview requests were formatted into broader structures of company and research life. My interlocutors were used to coffee dates with peers working at other companies, hour-long meetings with managers, and interviews with journalists. This last comparison occasionally posed problems, as my interviewees were wary about how public their words would become, or whether, reflecting a general suspicion of journalists, they would be used against them. So, although the topics addressed and the angles of questioning might be unfamiliar, these interviews were more of an extension of participant observation fieldwork than a distinct complement to it, a way for me to participate in my interlocutors' "interview culture" (Hockey 2002).

Screen work

Once inside a company, a more mundane problem presents itself: when everyone is working at computer screens, how is the ethnography supposed to know what is happening? The journalist and critic Quinn Norton has noted a problem that faces those who write about computer work:

> There is an aesthetic crisis in writing, which is this: how do we write emotionally of scenes involving computers? How do we make concrete, or at least reconstructable in the minds of our readers, the terrible, true passions that cross telephony lines? Right now my field must tackle describing a world where falling in love, going to war and filling out tax forms looks the same; it looks like typing. (Norton 2013)

This is a familiar challenge to anthropologists of work, who have had to reckon with field sites where "subjects' primary activities are speaking on the phone and typing on computer keyboards, leav[ing] little room for productive observation without conspicuously disturbing their work" (Dornfeld 1998, 23). Researchers have sought solutions in video recording of typers and their screens (e.g. Suchman 1996), a solution which causes horror in members of contemporary "tech" workplaces, where personal life bleeds into the workday and screens show not only the rudiments of work, but also those "terrible, true passions" cited by Norton. On a few occasions, I managed to talk my interlocutors into walking me through code they had written, or narrating to me a segment of their work, but people were so reticent to do so, and the resulting observations so removed from the "ordinary" work of coding, that it was much more productive to participate in conversations among co-workers, to listen in on meetings and casual banter about quotidian problems, and to talk about techniques face-to-face. I return to this problem of observing computer work and its significance for the ethnographic attention in the concluding chapter.

Corporate Heteroglossia

As evident in the beginning of this chapter, much talk in the world of music recommendation is organized around the actions of companies. It is common to hear phrases like “Spotify does this,” “Pandora says this,” or “Beats thinks that” from people working in the field. Particular approaches to recommendation are often associated with specific companies, such that “Pandora” can stand for human annotation of music, “Beats” can stand for expert curation, and “Spotify” can stand for large-scale collaborative filtering. This poses an interpretive challenge as these claims are not literally true. There is no good reason to presume that corporations—which are composed of many people, organized into many groups—will cognize, decide, or act in unison, in spite of corporate communications divisions whose job it is to construct a coherent, well-branded public-facing image. And although they come to stand for particular approaches to recommendation, all of these companies employ a variety of shared techniques beyond their stereotype method, in many cases even drawing on the same underlying data sources.

This internal variety sometimes manifests as heteroglossia in corporations’ public speech. For example, in Trent Reznor’s Beats Music manifesto, several voices are evident: the changes in register from the nostalgic saccharides of “falling in love every night” and music that “breathes and bleeds and feels” to oblique digs at competitors whose “code can’t hear” to the business-speak of “fun solutions” that “integrate.” Especially for nascent companies like Beats Music, messaging shimmers out of phase across platforms and incongruous actions reflect internal divides. It would be a mistake to take a tweet, composed by a college-aged social media intern, as an absolute statement of company positions, though Executives in public interviews, social media managers on Twitter or Facebook, reporters interpreting press releases, and employees talking off the record all constitute a messy discursive version of the corporation. These accounts are then supplemented by computer programs, which are themselves heteroglossic productions, with various teams responsible for their various parts, often only loosely coordinated with each other.

The challenge is to reckon with a ubiquitous way of speaking that presumes corporate actions to be intrinsically coherent, while recognizing that, in actuality, companies are variously anarchic sets of people and techniques that by no means have to act in concert. For outsiders to these companies, such a simplification is nearly unavoidable: the interiors of corporations can be as obscure as the interiors of others' minds, and to the extent that the interior is obscured, the actor in question appears unified. (This makes "the corporation" much like "the algorithm.") So, although I will continue to emphasize the interpretive and expressive heterogeneity of corporate action, I will also continue to use company names as my interlocutors do: as shorthand for messier sets of actors. Sometimes—as when interpreting a baffling business move—it is important to remember that companies have interiors. Other times—as when listening to an engineer describe their competitors—it is important to recognize that corporate identities are cognitively salient shorthand.

These challenges in specifying corporate identity are not unlike those that face attempts to define "algorithms." People often talk about algorithms as though they might be investigated directly, if only they had access to them and the expertise to interpret them. In practice, however, this clarity is evasive: ask an engineer to show you the algorithm, and they will give you a puzzled look. This is in part because there are usually multiple things that might be called “the algorithm” working in concert—one system profiles listeners, and its outputs are matched to another system that profiles music, and the matches, calculated by another system, are then put into a playlist by yet another. Specify more narrowly—“Show me the shuffling algorithm that sorts songs”—and you’ll hear that it is not all that interesting in isolation.[6] As one engineer on the recommendation team at a large music streaming company told me, it wouldn’t be that big of a revelation to release all the “algorithms” his company used for things like personalization—they wouldn’t say much about how the product works because they are so closely tied to the peculiarities of the data they operate on. It is that data, the collection of tracks and metadata, user listening history, and the like, that is really valuable. Algorithmic systems, as I am arguing, are heterogeneous things whose component parts seem to extend beyond themselves.

***

In the next chapter, I delve into the historical and conceptual origins of algorithmic recommender systems, as technologies for mediating between individuals and overwhelmingly large archives. Recommender systems can be understood as an example

[6] Or, as in a blog post from a Spotify engineer detailing a new shuffling algorithm (https://labs.spotify.com/2014/02/28/how-to-shuffle-songs/), you’ll learn the abstraction that guided its development, but you won’t learn if it actually works this way in practice (either because of peculiarities of the data it works on, errors in implementation, or unexpected interactions with other algorithmic parts of the system) and you won’t learn *how* it works in practice—both in the relatively objective sense of which outputs it generates and in the more subjective sense of how shuffled it *feels*. (On shuffle in particular, see Powers 2014.)

of “preferential technics”—systems that draw the anticipation of taste into circulatory infrastructures. The variety of recommender techniques reflects a variety of theories about taste, embedded in technical arrangements that are not alien to or simply context for taste, but are rather integral to its operations. Thinking of taste itself as kind of a collective technique mediating relations between persons and large sets of things, the apparent paradox of algorithmic recommendation—its accounting for taste—seems to resolve: taste and technology are not opposed as “subjective” and “objective,” but are rather fundamentally related in the interaction of techniques. The following three chapters build on this foundation, exploring how taste and technique interrelate in three key areas of music recommendation: listeners, music, and listening.

In chapter 2, I examine how the developers of commercial music recommenders understand their listeners. The purpose of algorithmic recommendation has come to be understood as “hooking” listeners—capturing their attention so that they continue to listen. Drawing on work in the anthropology of traps, I suggest that it is useful to think of recommender algorithms as “captivating technologies”: like other traps, their design is not informed by uniquely “functional” concerns that can be separated from arbitrary “cultural” variation. Rather, their design is informed by trappers’ understanding of the entities to be trapped, including broader cultural imaginaries. Among my interlocutors, music listeners were understood to vary primarily in their avidity—their enthusiasm for music and consequent willingness to interact. As a result, recommender systems were optimized to capture the attention of the musically “indifferent.”

Among the developers of music recommender systems, music is talked about as occupying or constituting a "space" in which listeners and algorithms travel. In chapter 3, on music, I investigate a common understanding of this space as a kind of landscape to which developers tend, identifying themselves as "park rangers" or "gardeners." Where extreme positions might hold that the music space is either discovered or invented, this pastoral imaginary cuts a middle path, placing the work of tending to musical space at the intersection of the natural, the cultural, and the technical. Although popular discourses about the internet and large databases consider them to be "post-geographic," allowing for novel or latent modes of connection that transcend physical proximity to emerge, they are "re-territorialized" by both this pastoral relationship to space and the persistence of geographically-bound understandings of cultural variation among those who tend to it. In other words, it is common to encounter genres with names like "Thai Hip Hop" or to find people who understand the variability of taste, space, and "culture" in national terms.

In chapter 4, I pursue the understandings of listening that inform the algorithmic processing of musical sound. Anthropological approaches to sound have emphasized embodiment, affect, and presence — features that seem to be excised in processes of quantification. However, as I witnessed among my interlocutors, the equation of sound and number is not simply a rationalizing process that leaves number untouched. Rather, people have developed a variety of listening practices for hearing their numbers, and these listening practices are a persistent mode of evaluation in day-to-day work: researchers sonify their high-level summaries of audio data, and they learn to listen to their recommender algorithms' output, drawing their own perception into the loop.

In the conclusion, I draw out parallels between mediated listening practices and the work of the ethnographer in contemporary fieldwork. Like recommender algorithms, ethnography is a set of techniques for directing attention. Attention's contemporary mediations offer a setting in which to consider how ethnography comports with other modes of paying attention (in particular those premised on "immersion" or other encounters with scale), amidst the emergence of a popularly recognized "attention economy."

CHAPTER 1:

PREFERENTIAL TECHNICS

The Celestial Jukebox Brought Down to Earth

It is not hyperbole to note that a revolution has occurred in the way that we as a society distribute data and information. (Fields 2011, 3)

Music like other online media is undergoing an information explosion. (Anglade 2014, 15)

This rapid increase in the quantity of available songs can lead to severe information overload. *Paradoxically, the overabundance of content can make it more difficult for a user to decide what to listen to next.* (McFee 2012, 1)

The three PhD dissertations quoted above were filed during the course of my research into algorithmic music recommendation. They begin—like countless other blog posts, newspaper articles, and scientific papers on music recommenders — by noting that listeners today have access to music at unprecedented scale. This is a long-awaited technical condition in the music industry, which at some point in the early 1990s was named "the celestial jukebox": a system that would make all music available to anyone, at any time, anywhere (Burkart and McCourt 1999).[7]

For the authors of these dissertations, the celestial jukebox was embodied in on-demand streaming services. These companies host large catalogs of music, their rights privately negotiated with record labels, and their subscribers can, for a fee or by listening to

[7] The term "celestial jukebox" makes some of its earliest appearances in Paul Goldstein's 1994 book *Copyright's Highway: The Law and Lore of Copyright from Gutenberg to the Celestial Jukebox* and in a 1995 Clinton administration report on intellectual property and information infrastructure, where it is an object of worry about copyright's inaptitude for such freely flowing cultural material.

advertisements, stream any song to their computer or smartphone. After a decade of halting starts, during which companies struggled to obtain the necessary deals with record labels, sustainable subscriber bases, and technical infrastructures to support streaming's data requirements, a new generation of services launched in the US around 2010. In their marketing materials, these companies eagerly claimed the qualities of the celestial jukebox: Spotify advertised "All the music, all the time"; Rdio, "Unlimited Music Everywhere"; Deezer, "Listen to all the music you love, anytime, anywhere"; MOG, "All the Music You Want." Rhapsody, one of the longer-lived companies in this category, declared itself "THE Celestial Jukebox!" as early as 2001. My fieldwork from 2011–2015 caught music streaming services in a period of dramatic growth: from 2013 to 2014, the number of on-demand music streams grew 60.5%, and by 2014, 67% of US listeners listened to music online (Nielsen 2014). It seems as though music services have, as one music technology journalist put it, at last reached the "promised land" (Bylin 2014).

But the arrival of the celestial jukebox, if this is it, has been less heavenly than expected. Critics note that the jukebox's defining superlatives—all the music, all the time, anywhere—have not actually been achieved: not all music makes it into the catalog (not even all *recorded* music does) and access is not universal. Although my interlocutors recognized these problems, they saw them as issues that would resolve over time, with the spread of technical infrastructure and the growth of streaming as the standard mode of music listening. Instead, they focused on what they saw as a problem endemic to large-catalog streaming services, which would only get worse as they grew: faced with so many options, people find themselves overwhelmed. As one of my industry interlocutors put it:

> With 30 million songs, how do you even begin? You begin, of course, by saying, "Hey, all the music in the world! Cool! I can listen to that Dave Matthews Band album I had on CD but never unpacked after my last move!"
>
> Then, 40 minutes later, it's over, and there are 29,999,988 songs left.

This problem is known by psychologists as information or choice overload: "although the provision of extensive choices may sometimes still be seen as initially desirable, it may also prove unexpectedly demotivating in the end" (Iyengar 2000, 996; see Schwartz 2004; see also Andrejevic 2013).

In the case of the "celestial jukebox," this problem manifested as listeners bewildered by large catalogs reverted to a small subset of music that they could easily draw to mind. As a result, they wouldn't explore the breadth of music available to them, they wouldn't discover the less popular artists in the "long tail," and, in the view of my interlocutors, the celestial jukebox's potential is wasted. For the companies that aspire to celestial jukebox status, this poses a business problem: increasing the size of the catalog may make people *less* likely to listen to it, turning instead to services like terrestrial radio, which require fewer choices, or to personal collections of music, which are more familiar, materially constrained, and easier to navigate. For my interlocutors, however, this is more than a business problem—it is a calling: if the celestial jukebox were realized, it could be horizon-broadening for listeners and career-making for musicians, but only if listeners venture beyond the music they already know.[8]

Introducing listeners to music that they don't know about, but might enjoy, is known as "music discovery," and it may be the closest thing to a universally shared value among the diverse parties involved in music recommendation. Academic researchers like the ones cited above often take music discovery as a self-evident goal. Record companies want listeners to find more of their products to consume, music services want listeners to keep using their platforms, and the engineers of recommender infrastructures see the discovery of new music as a worthwhile end in itself. As one head of engineering put it at an industry panel on music discovery in San Francisco: "we owe it to people to help them find music they love."

Music recommenders are built as aids to discovery, to provide navigational help to listeners who might like many things in the catalog, but wouldn't like everything. They mediate between the individual, with limited horizons, and the largeness of the catalog, which exceeds it. This tension animates the work of recommender systems over time: though they rely on profiles of user taste and are sometimes critiqued for constraining users inside these profiles, they are intended to operate at the knife-edge between the already known and the new. Although recommender systems are sometimes criticized for imposing biased contours onto the celestial jukebox and reifying their users (as in

[8] This idealism is not limited to the employees of technology companies. The dream of the celestial jukebox, fully realized, resonates with the ethnomusicologist Alan Lomax's late-career vision in the 1990s for a "global jukebox" of the world's music, which would

Alex Galloway's take on recommendation as "a process of interpellation" in a "hegemonic pattern" [2004, 114; see also Burkart and McCourt 2006]), their builders see them as potentially liberatory, in the service of expanding their users' musical worlds (not isolating them in homogeneous "filter bubbles") and realizing the promise of digital distribution for less popular artists in the "long tail" (Celma 2010).[9]

In this chapter, I delve into the relationship between overload, taste, and technology. Overload, I suggest, is not a consequence of particular technologies, but rather an enduring structure of feeling about the difference in scale between individuals and archives. Recommender systems are understood by the people who make them as technologies for managing this scalar mismatch. They do so by modeling and anticipating user tastes, drawing these anticipations into the technical infrastructure of music circulation. I call this interrelation of technical infrastructure and theories of taste "preferential technics." Preferential technical systems have been objects of critique, as noted in the previous chapter, for their blending of the supposedly distinct domains of taste and technology. Here, I outline an anthropological approach to studying preferential technics, rooted in parallel traditions from the anthropology of art and the anthropology of technology. These traditions suggest that, counter to popular common sense, taste and technology are not opposites, but are rather interdependent *techniques*

---

allow listeners to hear music from any culture around the world, again thanks to computers (Szwed 2010).

[9] This was also the sentiment behind a third-party application called *Forgotify*, which would randomly play tracks from the Spotify catalog which had never been played before. Spotify declined to comment when the app went viral, noting that the company provided its own discovery tools. Likely this reticence to comment was due to the sobering statistic that drove *Forgotify*'s press coverage: 20% of the songs on Spotify, some 4 million of them, had never been played (Palermino 2014).

for interacting with the broader world. Consequently, preferential technics can be understood as the mingling of techniques that are ordinarily segregated into two separate domains.

Typical critiques of these systems presume a distinction between taste and technology: these systems fail either because they try to blend domains that are inherently separate, or because they fail to make technical functionality align with cultural logics. If taste and technology are both understood as matters of technique, then this oppositional understanding comes undone. Continuing to draw on the anthropologies of art and technology, I suggest that the anthropological study of preferential technics should not concern itself with whether these techniques work (or whether the theories they embody are correct), but rather should investigate how they make sense to the people who pursue them. These techniques do not need to be socioculturally correct (according to legitimated theories from the social sciences) in order to persist, and they can constitute "performative infrastructures" (Thrift 2005, 224), which bring the world into alignment with the theories they embody. In this view, the role of anthropological theory is not to explain or correct theories and techniques found in the field, but rather to provide new ways to pay attention to them. Drawing on work that proposes a sideways (rather than hierarchical) relationship between anthropological theory and theory in the field (e.g. Maurer 2005; Riles 2011), I propose a "resonant anthropology" that sounds out similarities between theories and practices found in the field and in the discipline. This approach requires taking anthropology itself as a potential object of analysis, while being open to the local plausibility of theories found in the field.

Overload as Doxa

One of the leading academic researchers of recommender systems began a 2012 webinar in a strange way: "I'd like to start back about 20,000 years ago," he said, with a slide titled "Ants, Cavemen, and Early Recommender Systems—the emergence of critics." He was beginning, it turned out, like the dissertations cited above, by talking about the situation of overload in which humans find themselves. But unlike those dissertations, which focused on the contemporary problems of the celestial jukebox, he took a longer historical scale. Prehistoric humans, he speculated, had evolved to take advantage of each other's knowledge as a technique for interacting with the world. If a caveman came upon a bright red fruit, he did not have to eat it to know if it was poisonous—he could ask other cavemen. He did not require direct experience of all possible options, but instead relied on others. Those who insisted on tasting every possible fruit would not only find themselves overwhelmed by the task; they would find themselves removed from the gene pool. Eventually, he argued, this role was professionalized by scientists and engineers, and in the cultural domain by the role of the critic. "We are all pretty familiar with the idea that we take the opinions of others into account," he said. Recommender systems were not a break with earlier ways of interacting with the world, but rather the latest in a species-old sequence of techniques for sorting out what people wanted (and what they didn't) from the often bewildering vastness of the world.

For the people this dissertation is about, the idea that human life is characterized by encounters between people with limited horizons and sets of objects that greatly exceed their knowledge is doxic in Pierre Bourdieu's sense: a self-evident frame for action, difficult to bring into critical focus or to reject (Bourdieu 1972). The invocation of scale was a standard part of recommender systems developers' description of their own work: at a weekend hackathon, a coder told me about the incredible amount of music being made today thanks to home computers; a product manager explained to me that democratized access to music production meant that more music than ever landed "in your inbox"; an engineer marveled at how much more music from around the world was now available to ordinary listeners. Writers of blog posts, magazine articles, press releases, and dissertations (like those cited above and this one) seemed inexorably drawn to introduce their work by first noting what had come to seem obvious: contemporary music listeners have access to more music than they could ever hope to listen to, and they need something to help them manage it. To even explain what a recommender system is requires first introducing the idea of overload, which simultaneously seems like a new problem, entangled with advanced in information technology, and an old one, a basic condition of existence in a world with horizons beyond one's own.

The term "information overload" was popularized by the futurist consultant Alvin Toffler in the 1970s, following the publication of his *Future Shock* (which would eventually pick up the evocative subtitle "a study of mass bewilderment"). As demands on their attention grow, he wrote, people "may well find their ability to think and act clearly impaired by the waves of information crashing into their senses" (Toffler:1970, 354). In the same year, Stanley Milgram published an article on "The Experience of Living in Cities," describing how city dwellers adapted to the overwhelming stimulation of urban life —"urban overload"—by narrowing their focus and ignoring the needs of others (Milgram 1970). This line of thought—that people can be overloaded with information and stimulation—was adapted from system analysis, and it was aggressively pursued in the management and computing literature (see reviews in Eppler and Mengis 2004; Edmunds and Morris 2000; e.g. Whittaker and Sidner 1996), as it pointed to a potential drawback of the "informatized" workplace: more information does not necessarily lead to better decisions.

These problems would come to be typically associated not with cities, but with information technologies. Recently, the historian Ann Blair has recounted early modern experiences (ca. 1550-1700) of information overload as scholars faced a "multitude of books" and worried about their "confusing and harmful abundance" (Blair 2003, 11; see also Blair 2010; Ellison 2006}. Even earlier worries about "the abundance of books" can be found in the first century AD work of Roman philosopher Seneca. By 1982, the president of the Association for Computing Machinery was already complaining about the quantities of email he received:

> In my own situation, which is not unique, I must deal with a constant barrage of information. [...] Beyond the riptide of normal business mail lies a tidal wave of electronic junk mail [...] In our discipline we are liable to choke on our effusion if we do not effectively address the growing problem of the quantity of information we produce. (Denning 1982, 164, 165)

By the 1990s, popular and academic critics were identifying something they called the "attention economy": if economies concerned the distribution of scarce resources, then the so-called "information economy"—characterized by the abundance and availability of digital information—was really an "attention economy," since it was consumers' attention that was the *limited* resource (e.g. Goldhaber 1997). 24 years after his first letter on email, Peter Denning suggested that his warning had been apt, and "the tsunami arrived," and "technology is the source of these afflictions" (Denning 2006, 15, 16).

During my fieldwork, the face of overload was the on-demand streaming service, but before then it was located elsewhere. A dissertation on music recommendation completed in 2008 cites the growth of personal MP3 collections: "Personal music collections have grown, aided by technological improvements in networks, storage, portability of devices and Internet services" (Celma 2010, 9). A decade earlier, one of the first theses on music recommendation—and on algorithmic recommenders more generally — was a 1994 master's thesis from the MIT Media Lab, which set the growing availability of music (on compact disc) in the context of other products:

> Recent years have seen the explosive growth of the sheer volume of everyday things. The number of products, books, music, movies, news, advertisements, and the flow of information in general, is staggering. This truly is an "information age." The volume of things is considerably more than any person can digest. A person could not possibly filter through every item in order to select the ones that he or she truly wants and needs. (Shardanand 1994, 13)

The thesis went on to list how many CDs were available through the mail-order BMG Compact Disc Club (over 14,000), how many videos were available to rent in the average Blockbuster video store (6,000), and how many books were in the Library of Congress (15,700,905).

This feeling of overload, particularly as tied to the emergence of the World Wide Web, was the animating force behind the emergence of recommender systems research as a field. In the early 1990s, computing researchers began to seek new techniques to aid users in finding information in growing databases of things like emails, newsgroup postings, and music recordings. A 1992 issue of the Association for Computing Machinery's *Communications* was dedicated to the growing field of "Information Filtering," which was becoming vital "to control the potentially unlimited flux of information" available online (Loeb and Terry 1992, 27).

To solve problems like this, these researchers explored a variety of techniques for sorting through data: giving users finely tuned search tools, modeling their interests, extracting information about the content of items in the database, and so on. But what kicked off recommender systems research as a distinctive field was the advent of a technique called "collaborative filtering," developed by researchers at Xerox PARC for sorting email. They called their system "Tapestry": rather than conceiving of the user alone, up against the massive database with only her search tools to help her, collaborative filtering architectures linked users together, so that one person's filters could be shared with another, collectively "weaving an information Tapestry" (Goldberg et al. 1992).

Tapestry required users to actively sharing their filters with each other. Soon thereafter researchers at the University of Minnesota and MIT developed systems (called GroupLens and Ringo) to automate this process, matching users with each other on the basis of their shared interests or, as one group put it, "automating 'word of mouth'"

(Resnick et al. 1994, Shardanand and Maes 1995). Derivatives of this "automated collaborative filtering" are the dominant form of recommendation today, partially powering the recommender systems of large companies like Amazon, Netflix, and Spotify. Today, "collaborative filtering" usually refers to this automated variant, which works at much larger scales than its predecessor. As some of its developers reflected: "You did not need to know the identity of those you correlated with to gain the benefit of their recommendations, unlike Tapestry where the benefits came directly from your personal relationships with recommenders" (Borchers et al. 1998).

The research community that grew up around these systems eventually started its own conference in 2007: RecSys, sponsored by the US Association for Computing Machinery. The conference draws researchers from industry and academia to share work with each other. I attended RecSys twice during my fieldwork, at its meetings in Dublin and Silicon Valley. Although the community has its roots in the collaborative filtering systems developed at the University of Minnesota, RecSys today hosts a variety of techniques for making recommendations in a wide range of domains. These techniques, as they draw on various types of data and use various techniques to correlate them, are informed by a range of theories about taste and the inabilities of existing systems to account for them. They remain united, however, in their reliance on the idea of information overload as a contemporary problem and algorithmic systems as a solution to that problem.

But what about the cavemen from the webinar? Though typically invoked with reference to new information technologies, the doxic notion of information overload is much more expansive than that. For my interlocutors it is, I propose, a generic way of understanding how people relate to the world, as relevant for putative "cavemen" as it is for contemporary "millennials." Take, for example, the typical response to arguments against the "filter bubble"—the idea that personalization technologies such as recommender systems isolate users in self-reinforcing information bubbles, making it harder to find contradictory opinions or to engage with people unlike oneself (Pariser 2011). Many of my interlocutors argued that the supposedly unmediated alternative to the filter bubble was no better: people routinely isolated themselves through ordinary homophily and geographical parochialism. This, they suggested, was also a filter, and all the more insidious for its apparent naturalness. In comparison to the potentially extreme and biased filtering effected by ordinary conditions of life on earth, algorithmic recommendation could be horizon-broadening. In the overload frame, everything can be understood as a kind of filtering, and there is no "natural" state absent mediation; in this, my interlocutors were in line with a substantial body of media theory literature which similarly argues for the omnipresence of mediating, filtering, translating practices (e.g. Bowker 2010; Peters 2015).

Though information (or choice) overload is typically imagined to be the result of specific technical conditions—on-demand streaming, personal mp3 libraries, mail-order CD clubs—its persistence over time suggests that it is neither new nor tied to specific technologies. Given this historical depth, information overload can be more precisely identified as an enduring structure of feeling (Williams 1977, 128–135): an anxiety about the difference in scale between individuals and archives.[10] For my interlocutors, although this anxiety manifested in relation to particular technical practices, this scalar mismatch was not simply a consequence of new information technologies but a fact about how people relate to the world. People inevitably attend to certain things and not others, relying on criteria that can be explicit or tacit, but which always exist in some form. We might imagine alternatives to overload, where the knowledge that the world exceeds one's horizons does not inspire wanderlust or anxiety (or where the notion of the world and horizons is displaced by something else entirely), but this is not the frame in which my interlocutors live and work.

Preferential Technics

The experience of musical choice overload is currently facilitated by what media historian Jonathan Sterne calls "the worldwide proliferation of mp3 files" (Sterne 2012, 188). On-demand streaming services typically provide music in the mp3 format—a digital standard for audio data. Mp3 stands for layer 3 of the MPEG-1 standard, and MPEG stands for the Motion Picture Experts Group ("a consortium of engineers and others formed with the support of the International Standards Organization (ISO) and the International Electrotechnical Commission" [Sterne 2006, 829]). In his exhaustive history, Sterne describes how, late in the 20th century, a set of electronics and recording companies came together with international standards bodies to produce, somewhat inadvertently, a technology that made it easy to digitally distribute and store music and which would eventually be blamed for throwing the music industry into crisis.

[10] One could call this experience of archival scale a "data sublime," the historical resonance of which will become clearer in the chapter on data landscapes. (See, e.g.

The mp3 is a standard for compressing data, but what makes it notable is that it is "lossy." Psychoacoustics research has demonstrated that human hearing is partial. The mp3 draws on this research to selectively omit sound that would not be heard by an ideal-typical human listener. For example, louder sounds "mask" or hide nearby sounds in similar frequency ranges, so such occluded sounds can be left out of the recording altogether. The production of the mp3 standard, as Sterne outlines, involved a variety of laboratory listening experiments that tried to formalize and measure phenomena such as "annoyingness"—at what levels did the slight distortions of the mp3 format become apparent and irritating? The mp3 was thus a technical infrastructure dependent on theories about how the humans at the output end of the system functioned,[11] braiding together the concerns of engineers, standards bureaucracies, and psychoacoustics researchers to produce a format that was designed for easy transmission, storage in bulk, and casual listening. These affordances facilitated, Sterne argues, the widespread distribution of illicit mp3 files through file sharing services like Napster at the turn of the century. A decade later, they are also supporting the technical infrastructure of on-demand streaming services.

Sterne argues that the mp3, in its anticipation of a certain kind of listening subject, is a form of "perceptual technics"—the intermingling of scientific models of perception, technologies of circulation and distribution, and the economic interests of corporations.

Davies 2015, who builds on Stallabrass 2007, or Mosco 2004 on the "digital sublime.")

[11] If elements of the listening and distribution assemblage changed, then the special accommodations of the format became more obvious: outside of casual listening

Mp3s anticipate listeners who will finish the job of transforming audio data into sound in a predictable way, and this listening labor produces, as it were, a kind of surplus: providers of mp3s find more space on their hard drives, more bandwidth in their transmission channels, and they can capitalize on this surplus by paying for fewer resources or packing more files into a smaller space. Perceptual technics integrates listeners' hearing into the infrastructure of distribution and circulation and capitalizes on it. "One might even say," Sterne writes, "that the mp3 is a celebration of the limits of auditory perception" (Sterne 2006, 828). Rather than thinking of the mp3 as simply another stage in a progression of sound reproducing technologies, Sterne points to the particularities and contingencies of the format, entangled with certain ways of understanding hearing and transmission.

We can understand recommender systems, by analogy with the mp3s that some of them filter, as a kind of *preferential technics*. Where the mp3 capitalizes on the anticipation of patterns in human hearing, recommender systems capitalize on the anticipation of patterns in human preference. Perceptual technics is "lodged between mechanics and biology," as Sterne quotes Bernard Stiegler (1998, 2; Sterne 2012, 53), and we might say that preferential technics is lodged between mechanics and *culture*. Both blend together the technical exigencies of circulation with theories about the functioning of audiences. Both mediate between the individual (understood as limited) and a large body of information that exceeds individuals' capacities. Where the mp3 effects, as Sterne suggests, "a concordance of signals among computers, electrical components and

---

environments or for listeners outside the ideal-typical human (lizards, children, or computers) the losses of the encoder can become audible.

auditory nerves" (2006, 837), recommender systems aim to bring about a concordance of music listening services, algorithmic sorting, and human preferences. By "preferential technics," I refer to these anticipatory technical practices, which aim to draw the preferences of users into circulatory infrastructures. Although one might argue (like my interlocutors above) that all circulatory infrastructures filter, intentionally or inadvertently, preferential technics is distinguished by taking this filtering as an explicit object of research and engineering. Preferential technics is facilitated by the frame of overload: in that frame, taste is just another filtering practice, and technical infrastructures can be brought into accord with it, to anticipate it.

Preferential technics bears some similarity to the "choice architecture" advocated by legal scholar Cass Sunstein (Thaler and Sunstein 2008; Sunstein 2015): a libertarian paternalist structuring of decisions like buying insurance, saving money, or becoming organ donors, which draws on research in behavioral economics and psychology to "nudge" people into making healthy, socially beneficial, or otherwise desirable decisions. I address these debates about manipulation and behaviorist understanding of human activity in later chapters; however, my primary goal here is not to mount a critique but to understand how people building these systems make sense of the world, and in particular the relationship between taste and algorithmic filtering. Occasionally, my engineer interlocutors would express paternalistic goals, wishing to steer mainstream listeners to "better" or more obscure music, though they rarely felt that they could actively do something about it. Some critics (e.g. Harvey 2014, Striphas 2015, Andrejevic 2013) argue that these anticipatory infrastructures pre-empt the operation of taste altogether—can one even have taste when a computer has anticipated one's choices? But in this frame, where there is no fundamental ontological difference between filtering effected by taste or by algorithms, the concern that algorithms might displace taste with something altogether different is not intelligible.

The key question for preferential technics, as it was for perceptual technics, is: *How are listeners anticipated?* These anticipations are built into infrastructures of circulation (as the psychoacoustic model of hearing is built into the mp3 format), and they tangle together the ostensibly separate sound and listener, through theories about taste, listening, and music. Like perceptual technics, preferential technics involves a messy coexistence of scientific research, ad hoc theorizing by engineers, contested standards, and — last but not least—commercial interest in the contouring of digital commodity flows. The development of preferential technics is thus situated at the confluence of many different streams of activity, from music industry strategizing to academic research in information science to informal "hacking" by musically-inclined programmers. Unlike the mp3, which has by now established firm dominance in the world of perceptual technics, preferential technics is still in flux, implemented in diverse ways. Through different models of listeners, different technologies for circulation and listening, or different business models, preferential technics varies. The following chapters document my efforts at capturing an emerging preferential technics in formation, while its central terms were still being negotiated, in the construction of new algorithmic systems.

Taste and Technology

My interest in the practices I call "preferential technics" stemmed from their apparently contradictory nature. Taste and technics are commonly assumed to be distinct domains, as separate as the human and technology from the previous chapter. In popular common sense, the apocryphal Latin epigram *De gustibus non est disputandum* reigns: there is no accounting for taste. Taste is privately felt, unavailable to reason, and the pinnacle of subjectivity. In this view, tastes cannot be explained, and consequently should not be argued about. Technology is the opposite: rationalizing, motivated by necessity rather than arbitrary election. Curiously enough, technology is also supposed to be unfriendly to argument, but because it is too *objective*. Where taste is essentially human, technology is essentially machine. When companies try to blend the two domains together, they receive critiques like this one leveled against Netflix by the journalist Felix Salmon: they do not (cannot) work, because they attempt "the systematization of the ineffable" (Salmon 2014).

This popular common sense falls apart in the face of everyday life. Not only do people argue about taste, as Steven Shapin suggests, "we argue about little else" (Shapin 2012, 176), populating a lively public discourse about the merits of various cultural objects. Kant called this paradox—though tastes are subjectively felt, people are nonetheless adept at coming up with explanations for them—the "antinomy of taste" (Kant 1914[1790]). And not only do people argue about taste, but a wide range of taste technologies (Shapin calls them "sciences of subjectivity") exist to help people develop and account for their tastes, from the various apparatuses of wine tasting to ratings platforms to the preferential technical systems discussed here: "If there is no accounting for tastes, that's news to the accountants" (Shapin 2012, 179).

The academic common sense about taste is a bit more complex: it can be found in the opening chapter of Pierre Bourdieu's goliath *Distinction*, on "The Aristocracy of Taste" (Bourdieu 1984). Taste, in this view, is a patterning of preference that emerges from and contributes to the structure of the social field. In other words, it is a tool for organizing and enforcing social distinctions. Plainly: fancy people have fancy tastes, while simple people have simple tastes (sociologists call this the "homology thesis"). While taste is arbitrary with regard to its objects (we could live in a world where opera is low class and heavy metal is high class), the structure of tastes is not arbitrary in itself. It manifests power relations and is a tool of symbolic exclusion (Bryson 1996). More recent work has complicated the homology thesis, noting the emergence of "omnivorousness" as a signifier of high status (Peterson and Kern 1996), but the basic sociological understanding remains the same. Taste is a function of social position, not of the aesthetic content of cultural objects. As Bourdieu put it: "Taste classifies, and it classifies the classifier" (Bourdieu 1984, 8).

These common senses lead to two common responses to the work of preferential technics. The first presumes that engineers represent the technological side of a technology-culture dichotomy and thus misunderstand cultural phenomena like taste. Their attempts at making sense of taste force it into a technocapitalistic Procrustean bed, corrupted by business interests and the engineering mindset. Engineers, in this view, are socially and culturally incompetent, having replaced any tasteful sensibilities they might have had with technological requirements. This attitude is summed up in a quip from the developer Maciej Ceglowski: "Asking computer nerds to design social software is like hiring a Mormon bartender" (Ceglowski 2011). The second common response takes the problem not to be the mingling of taste and technology, but rather an ignorance about how taste “really” works. In this view, if engineers learned their sociology, then they might be able to build a system that works in accord with legitimated social theory. This position is popular among social scientists at conference receptions and happy hours, who speculate on how they could fix Netflix or Pandora, if only people with technical chops would listen.

My argument is that both the popular and academic common senses outlined here are inadequate for understanding preferential technics. If there are problems with how engineers approach cultural phenomena like taste, these are not because engineers are ignorant or incorrect about culture, but because their work *constitutes* culture. There is not an intrinsic difference between cultural and technological domains, and the development of preferential technical systems troubles simplistic ideas about technology and culture alike. The supposed division between culture and technology becomes important in this view not because it is true, but because it is a dominant cultural frame in the world where my interlocutors work. Consequently, the builders of preferential technical systems have to reckon with this popular idea while pursuing work that seems to undermine it.

Rather than working from a critique in which social scientists hold social explanatory power and engineers hold technical explanatory power, I maintain that any definition of culture adequate to the world should take into account understandings of culture that exist in that world (Kroeber and Kluckhohn 1952; Strathern 1995; Fischer 2009). Ideas about taste do not need to be correct (nor explicit) to be built into infrastructures, and their success does not necessarily depend on their correctness, either. This has consequences not only for how to talk about preferential technical systems, but also for our understanding of theories about taste. In the balance of this chapter, I outline anthropological approaches to thinking about taste and technology in terms of *techniques*, which help make legible their similarity. I conclude by discussing what this approach to preferential technics entails for the theorizing of the social sciences in relation to the cultural work found in the field.

Techniques

On the very first page of *Distinction*, Bourdieu argues that "one cannot fully understand cultural practices unless 'culture,' in the restricted, normative sense, is brought back into 'culture' in the anthropological sense" (Bourdieu 1984, 1). This relativizing argument should be familiar to anthropologists: instead of talking about culture (or taste) as something that only some people *have*, we take it to be something that everyone *does*. The work of the social scientist in this regard is not to adjudicate between high and low cultures, but rather to understand how culture is variously done, expanding our frame of reference from the narrow scope of "high" culture to the broader world of practices, signs, institutions, relations, and so on that is both the subject and object of cultural life. However, in spite of Bourdieu's introductory remarks, anthropologists have been notably absent from the literature on taste, which has remained the province of sociologists.[12] Heuristically, we might ask: Why is there not an anthropology of taste, and if there were, what would it be like?

The distribution of disciplinary interest in taste has shaped the evidence used to understand it, and as a consequence, the kind of thing "taste" has become. As it has been sociologized, taste is intimately connected to the structures of the Western culture industry (Fenster 1991). The common sense understanding of taste as a homologue of social structure relies on a field of cultural options from which one can choose arbitrarily. Someone buying a CD, for example, can choose between heavy metal and opera at will — the recordings themselves are alike in all but their contents and the habitus of their listeners. The patterns in the resulting choices can thus be interpreted not as the result of material necessity or some other force, but as an outcome of arbitrary cultural differentiation. (This is not to say that the symbolic exclusions effected by taste do not have material consequences, but rather that we would not explain someone buying a top-40 album by saying that they need it to feed their family or fix their toilet.)

In traditional anthropological field sites, these broad fields of cultural goods were apparently absent—to have a favorite musical artist in a small-scale society is not the same as having one in 1960s Paris. While this is no longer true of anthropologists' field sites, it has contoured anthropological attention such that "taste" remains under the purview of sociology. The closest we get to an anthropology of taste is the anthropology

[12] Excepting, of course, the gustatory work of anthropologists like Paul Stoller (1989) and on food and the senses.

of art and aesthetics—a subfield defined in large part by one question: Is "art" a meaningful cross-cultural category, or is it a Western imposition? (see, e.g., Coote and Shelton 1992; Marcus and Myers 1995). The anthropology of art provides two desiderata for an anthropology of taste: First, it should relativize taste, not only in the rejection of high/low distinctions, but also in considering whether the system of taste itself might be variable. Second, it should be open to the plausibility and pragmatics of local theories, which multiply and contradict each other.

We can find a useful model for this effort in a more recent set of sociological inquiries that investigate taste as a kind of *technique*, looking to the situations in which tastes operate rather than to the broader social structures they have been understood to reflect. This work, on the pragmatics of taste (e.g. Hennion 2007), or the pragmatics of valuation more generally (e.g. Antal 2015), emphasizes the importance of encounters to the production of taste.[13] In this view, the Bourdieusian common sense, in which taste is "an arbitrary election which has to be explained [...] by hidden social causes" (Hennion 2007, 98), leaves an explanatory hole: tastes may map to social status, but that does not tell us about how they are acquired. The Bourdieusian common sense about taste lends itself to large scale statistical analyses, the pragmatic approach to the ethnographic study of everyday life, in studies of opera fans (Benzecry 2011), wine tasters (Shapin 2012), or music enthusiasts (Hennion 2001).

[13] In cultural studies, Ben Highmore has explored resonant issues under the heading of "social aesthetics," which we might just call the sociology of taste were it not for his central concern being moments of affective encounter (Highmore 2010).

Antoine Hennion's work with music enthusiasts, or *amateurs* (the term in French lacks the "non-professional" meaning it has in English), offers some insights for the concerns of this dissertation. *Amateurs* spend a lot of time setting up the conditions of their taste—they make listening rooms and prepare themselves for concerts, all to experience an encounter with cultural objects. Taste, Hennion suggests, is "a collective, reflexive, instrumented activity" (2007, 109). He analogizes these instrumented encounters to the movement of a rock climber on the face of a mountain: "The body is revealed in gesture and appears to itself, from whence comes pleasure" (Hennion 2007, 100). Taste is not latent in the climber and her social position (as the "sociological" explanation might have it), nor is it simply a feature of the rock (as the "aesthetic" explanation might suggest). Rather, taste is something like the interaction between climber and rock, an emergent, relational property that reveals capacities and differences hidden in both. It is not the high-level structure of difference that might be constellated onto a plot, but the grain of interaction which constitutes taste. This interaction, Hennion argues, "closely depends on its situations and material devices: time and space frame, tools, circumstances, rules, ways of doing things. It involves a meticulous temporal organization, collective arrangements, objects and instruments of all kinds, and a wide range of techniques to manage all that" (Hennion 2004, 6). Taste is constituted by ways of interacting with the world, and these techniques are supported and shaped by a wide range of other techniques—the varieties of human action that come together to constitute an opera hall, a compact disc, a guitar performance, or turntablism.

Research in the pragmatics of taste typically casts the Bourdieusian common sense (specifically the homology thesis, where taste is a function of status) as an antagonist, for its emphasis on large-scale social structure over a consideration of how taste is acquired and experienced. Omar Lizardo has argued that this is a misunderstanding of Bourdieu's argument, which assumes that his contribution is basically what Veblen argued in 1899: that taste corresponds with class (Lizardo 2014; Veblen 1899). Rather than representing a macrosociological departure from Bourdieu's practice theory, Lizardo argues, *Distinction* contains an account of how taste is acquired as part of a person's habitus—it is simply hidden deep in the latter part of of the massive book's "sprawling," "odd structure" (Lizardo 2014, 337).

For Bourdieu, the concept of habitus provided an alternative to visions of people as either free-willed subjects or vessels for structure (see Sterne 2003, 376), though the Bourdieusian common sense tends toward the latter understanding, in which tastes are determined structurally. Discrepancies between the received and the historical Bourdieu are not especially important for my purposes here; it is useful, however, to note that research in the pragmatics of taste comports with the historical Bourdieu's understanding of taste as a feature of the habitus—of a person's acquired practical knowledge—neither externally determined nor the free exercise of some intrinsic latent desire.

Habitus was also the term used by Marcel Mauss to denote the collection of "techniques of the body" possessed by people as members of a society in his 1934 article by the same name (Mauss 1973). These techniques famously included practices like basketweaving, using a hammer, sitting, or swimming, identifying them as culturally specific forms of patterned action, learned and potentially variable. A technique, for Mauss, is an "ensemble of movements or actions" (Schlanger 2006, 149), and even technologies that appear more exterior to the body—like cars or computers—"are deeply tied to techniques of the body, to the ways in which people learn to use and relate to their own bodies" (Sterne 2003, 80). So although Bourdieu tended to avoid the study of technology, one of his key concepts—*habitus*—has deep ties to technical concerns: how people come to know and use their embodied practical knowledge through interactions with a world already populated by practical knowledges and their crystallization in material arrangements we might call "technology." Jonathan Sterne argues that "technologies are essentially subsets of habitus – they are organized forms of movement" (Sterne 2003, 370). Seen in this light, tastes and technologies come to appear quite similar—they are forms of patterned action that shape and are shaped by interactions with the world, and they rely on environments populated by other forms of patterned action. Hennion's analogy to the rock climber makes more sense in this conceptual lineage: rather than thinking of taste as merely the exercise of preference and rock climbing as the exercise of necessity, we can think of both as contingent and interactive kinds of technique.

Descending from scholars concerned with taste and resurfacing amidst scholars interested first in technology, we find Tim Ingold writing on "The Anthropology of Skill," attempting to overcome the division between art and technology. He observes: "art has been split from technology along the lines of an opposition between the mental and the material, and between semiotics and mechanics" (Ingold 2001, 19), noting that the words' Greek origins—*ars* and *tekhne*—were not as opposed as their descendants have come to seem, and can both be translated as "skill" (see also Boellstorff 2008 on *tekhne*'s capacity for world-making). Where Hennion had his rock climber, Ingold has the maker of string bags, whose skilled manipulation of materials is responsive and cannot be reduced to a mental plan or set of material constraints (Ingold 2001, 22–25). Skill, like taste, is constituted by interactions. Ingold's "skill" is much like Mauss's "practical reason" or Bourdieu's "practical knowledge," the result of interactions of entities and their environments. Ingold, like Hennion, places great significance on the interactional nature of this quality, which varies widely because it is so contextually specific.

Common senses about taste and technology fit together, with complementary ambivalences about choice and constraint: taste is commonly considered a matter of preference, while technology is a matter of necessity. It is a common social scientific move to reverse these descriptions, emphasizing the determinedness of taste and the flexibility of engineering.[14] But the interactive understanding of taste and technology, as kinds of technique thoroughly dependent on wider environments already constituted by other techniques, allows another route through the swinging pendulum of choice and constraint. What had appeared to be two opposed poles can be brought together by thinking of them together. The anthropology of taste I develop in this dissertation is thus also an anthropology of technique—techniques for drawing out qualities of sound and music, so that they can be brought into new relations to listeners, techniques for apprehending the activities of listeners and channeling their attentions, patterns of action and sensation and analysis.

[14] See, e.g. Lizardo: "Preferences and taste do not drive action, but are an optional, always dispensable, commentary on what we cannot help but do" (2014: 356); and (Lemmonier 1992) on "arbitrariness in technologies" or (Pinch and Bijker 1984) on technologies' "interpretive flexibility."

Anticipatory Infrastructures

This understanding of taste and technology not as opposites, but as interdependent styles of technique, poses a problem for common popular and academic ways of talking about recommender systems. Namely, it draws into question what it means to say that a given recommender "works." As I learned over years of explaining myself to interlocutors in the field and the academy, this evaluative concern dominates discussions about algorithmic recommendation. I am frequently asked for my professional opinion: Which system works the best? Why does this particular system work so poorly? What should these systems do if they want to work better?

After I introduced myself as an anthropologist to the room at a 2014 music hackathon in the San Francisco offices of GitHub, an engineer came up to me to ask for advice. He was building a prototype recommender system that would match users to music based on mood, and he wanted my anthropological opinion on how the weather might affect people's mood, so that he could add the local weather as a signal to his recommender. A pair of men from a music data startup, which had been aggressively collecting a wide range of data about music and musicians, asked me if I could help them understand their users better. At a music business panel in Hollywood in 2013, an aspiring music startup founder suggested that I was a "bias detector"—as an anthropologist and outsider I could help point out where the assumptions guiding a company were off the mark. Maybe I could shed light on why people *really* like the music they like, or maybe I was there to critique the amateur cultural theories and biases of engineers, who were often pursuing strongly held, idiosyncratic ideas about music's significance.

It is not surprising that an interest in recommender systems is presumed to be an interest in making them work better. After all, recommenders are surrounded on all sides by evaluative logics. Their inputs are ratings (or user behavior interpreted as ratings); their permutations are internally evaluated in constant comparison tests for how well they perform according to various metrics. When they are released into the world, they are reviewed by technology critics whose understanding of criticism is typically limited to judging how well they work. The businesses that develop these systems are caught up in evaluative logics as well, reliant on venture funding that enters them into an arena where companies live and die in a routinized series of evaluations: How rapidly is the userbase growing? How rapidly is the rate of growth itself growing? All of these questions are reduced to one: Does it work?

This question is commonly understood (among insiders and outsiders alike) as a matter of anticipation: a system can be said to work if it accurately anticipates its users. Among the developers of recommender systems, this idea has been formalized in a set of evaluation practices that I describe in more detail in the next chapter. Briefly, these evaluations compare the ratings or behavior that a system predicts to the ratings or behavior that eventually transpire. The more accurate these predictions are, the better the users have been anticipated and the better the system is understood to work.

I call this the "representational understanding" of preferential technics: systems work or fail roughly according to the adequacy of their internal representations of users. This understanding is implicit in many of the critiques of recommenders I have heard in the field and the academy: these systems should work better, and working better means working from a more adequate understanding of what taste is and how it works. When this argument comes from social scientists, it usually means that engineers should read more sociology; when it comes from engineers, it usually refers to some "obvious" common sense about taste that guides their own work, but not their competitors'. (If you ask the competitor, you will likely hear the same argument about common sense, but with the company names reversed.)

Technologies more generally can be understood as anticipatory: standardized bolt sizes anticipate specific wrench sizes, social media platforms anticipate specific kinds of use, and urban infrastructures anticipate specific kinds of bodies (and, for all, vice versa). But the critical feature of anticipation (and that which draws much critical attention) is that there is always a mismatch between what is anticipated and what arrives, from the minor to the severe: individual listeners do not all hear the same as the ideal typical listener constructed in the lab, users want to do something other than what a platform was designed for, or certain classes of users, ignored during design, are left unable to use at all.

Because absolutely correct anticipation is impossible in the flux of the world, ordinary experience is constituted by a shifting in and out of phase of expectations and actualities. Technology critic Sara Watson describes this uncanny effect in the context of online advertising as the production of a "data doppelgänger"—a not-quite reflection of the self, whose slight variations draw attention to anticipatory infrastructures (Watson 2014). In the case of classificatory infrastructures, Geof Bowker and Susan Leigh Star argue that anticipatory failures produce “torque”—the twisting of unanticipated lives and bodies into ill-fitting schemata (Bowker and Star 2000) For the feminist philosopher of music Robin James (2014), these mismatches between dominant organizing structures and excluded people produce “dissonance”—an out-of-syncness registered as displeasing. This out-of-syncness is differently experienced by different people relative to anticipatory technologies: those in dominant subject positions feel the *unheimlich* dis-ease of the uncanny valley, while those on the margins are torqued.

While these critics lay out the harms of anticipatory failures, it is tempting to presume that the mismatches they describe could be resolved by becoming better anticipators, closing the gap between predictions and eventualities. For recommender systems, this might mean more adequately representing users’ tastes, relying on more accurate theories about how those tastes are patterned and formed. This is the kind of input engineers would ask me for during hackathons (when they weren’t already sure that their own ideas about culture were accurate). For example, Christian Sandvig has argued against practices he calls “corrupt personalization,” in which preferential technical systems are distorted in the service of inappropriate goals, from payola to propaganda (Sandvig 2014). To solve this problem, a service simply needs to stop the distorting corruption and to serve what Sandvig calls the user’s “authentic interests” (Sandvig 2014).

But, as Sandvig notes, “it is impossible to clearly define an ‘authentic’ interest” (Sandvig 2014), and it is similarly impossible to close the gap between a “distorted” representation and the world. As I outline in the next chapter, this is something the developers of recommender systems have realized as their evaluations premised on prediction stalled out under what they call a "glass ceiling"—they have been moving toward evaluations that focus on recommendation not as a prediction, but an *interaction*. This interaction, playing out over time, is not a spot-the-differences game between two visual images, one anticipated and one actual. Rather, it is more like the time-domain unfolding of sound, where anticipatory frequencies, sounded with enough force, have the power to bring others into resonant vibration with them, but may also be subsumed by them. The "wrong" prediction is thus not a straightforward thing to identify in most cases, as errors manifest as play in the shuffling margins of human action and choice.

To understand taste and technology as techniques directs our attention not to representation but intervention (see Hacking 1983). If we take taste and technology to be intertwined rather than opposed, then we should look to technical systems not as mirrors of taste, but as active elements in the formation of tastes. They are engines, not cameras (MacKenzie 2006), "performative infrastructures" that can shape human activity (Thrift 2005, 224). This understanding foregrounds the significance of technical supports—the infrastructures where tastes are formed and held, as Hennion argues (2004). Theories of taste are embedded in these infrastructures not as pictures that may be more or less distorted, but as collections of choices and tendencies, not all of which are explicit. Taste is relational and thus depends on the environment in which relations can be made: to have taste amidst the mid-20th century music industry was different than to have taste before the advent of recorded music, and it may yet be different in the context of on-demand streaming services. Taste might be otherwise, and if taste is closely linked to its technical supports, then the emergence of new forms of preferential technics is a sign that taste may be in the process of becoming something else.

Methodological Philistinism

I found it difficult to escape the evaluative frame, and this posed some problems during fieldwork. Trying to explain what I was up to and what anthropology was for to my interlocutors (this, I have come to believe, is the underappreciated motor of the contemporary ethnographic encounter), I was often at a loss for words: How was I supposed to explain that I was interested in how recommender systems worked, but I did not care how *well* they worked, when evaluation was the whole raison d'être of recommendation? I found it distressing to be interpellated as a source for anthropological evaluations, when I thought the very idea was a contradiction: anthropologists, with their interpretive charity and cultural relativism, are notorious for *not* evaluating the people and practices they study. Nonetheless, when it came to recommender systems, I found many of my colleagues unusually willing to take on the evaluative role. Many of them had dabbled with ideas about how to build a good recommender system on a sound sociocultural foundation. If only engineers knew what taste was *really* like, then their systems would work. The dominance of the evaluative frame speaks to the cultural position of these technologies in the ordinary lives of social scientists: after reading an ethnographic account of horticulture in a small-scale society, no one replies, "Yes, but are the yams any good?"

Anthropological critique is, at its best, more complex than identifying something bad, foolish, or incomplete in the field (Marcus and Fischer 1986)—it is comparative, like

Margaret Mead using her field research on sex and temperament in Papua New Guinea to undermine American assumptions about gender and sexuality (Mead 1935); it is reframing, like Sharon Traweek or Hugh Gusterson approaching the practices of high-powered high-energy physicists as a species of ritual (Traweek 1988; Gusterson 1996). It does not work within given frames, but rather against them, sharing a goal with Alfred Gell's vision for the anthropology of art, "the ultimate aim of which must be the dissolution of art, in the same way that the dissolution of religion, politics, economics, kinship, and all other forms under which human experiences is presented to the socialized mind, must be the ultimate aim of anthropology in general" (Gell 1992, 41). Or, as Edward Sapir put it, the "destructive analysis of the familiar" (Sapir 1921, 94). In other words, the goal should not be to take the evaluative frame of a technology on its own terms, but rather to make it strange and in the process learn something about evaluative frames and technologies more generally. Drawing on the literature, we can find related arguments for ignoring whether our subject matter "works" from both the anthropology of art and technology.

Gell proposed that the anthropology of art must avow a "methodological philistinism" — "an attitude of resolute indifference towards the aesthetic value of works of art" (Gell 1992, 42)—if it ever hoped to escape the untenable assumption of universal aesthetic criteria and, consequently, to be able to reckon with art produced in cultures other than those that imagined the criteria. Gell modeled his philistinism on the sociologist Peter Berger's "methodological atheism" in the study of religion (Berger 1967): to study religion as a social phenomenon means that "theistic and mystical beliefs are subjected to sociological scrutiny on the assumption that they are not literally true" (Gell 1992, 41). Gell goes even further in suggesting that his philistinism follows from Berger's atheism, arguing that methodological atheists have simply displaced their faith in god into a faith in art: "we have sacralized art; art is really our religion" (Gell 192, 42). This anti-aesthetic move parallels Bourdieu's rejection of aesthetics in his analysis of taste in *Distinction*. Not only does the social analysis of taste not consider the aesthetic quality of objects as a significant factor in the formation of taste, it rejects those qualities as a starting premise. Social analysis, in this view, is predicated on the ignorance of what participants see as central: religion studied without belief, art studied without aesthetics. The combative tone—"philistinism," "atheism"—speaks to the dominance of those frames in the worlds of the analysts: to not rely on religious or aesthetic explanation in religious or aesthetic domains would be seen as atheism or philistinism, when it might more moderately be called "agnosticism" or simply "paying attention in a different way."

This approach is not limited to "soft" domains like religion and art. Social studies of technology have similar heuristics. In his 1992 review of the social anthropology of technology, Bryan Pfaffenberger suggested something similar: given that technical practices around the world are often dependent on supposedly non-technical "ritual" or "cultural" practices—to coordinate labor (Lansing 1991), to produce technicians with the appropriate worldview (Orr 1996), to organize resources (Rappaport 1968)—it is difficult, if not impossible, to distinguish the cultural from the technical ahead of time. Faced with the realization that all technical systems are *socio*technical systems, anthropologists of technology have a hard time separating the technical (i.e. effective) from the cultural (i.e. arbitrary or stylistic):

> I would therefore argue that the social anthropology of technology, against all common sense, should adopt a principle of absolute impartiality with respect to whether a given activity "works" (i.e. is "technical") or "doesn't work" (i.e. is "magico- religious"); only if we adopt such impartiality do the social dimensions of sociotechnical activity come to the fore. (Pfaffenberger 1992, 501)

Pfaffenberger adapts this premise from the sociology of science, where it is known as the “principle of symmetry” (Bloor 1976; Latour 1987). Although social causes are usually only invoked for false beliefs (e.g. Lysenkoism was popular in Soviet Russia because ideology trumped objectivity), the principle of symmetry follows from the Wittgensteinian insight that truth alone cannot compel belief (e.g. people did not start to believe in Mendelian genetics simply because it was correct). Again, social analysis sets aside from the outset distinctions that participants see as central: technology without efficacy, science without truth. To hold all the potentially relevant parts of a sociotechnical system in the frame, the anthropologist of technology must be open to the roles played by norms, artifacts, rituals, beliefs, etc. in addition to what interlocutors or other experts hold as technically salient.

Recommender systems, with their entanglement of taste and technology, are caught in this pincer movement of the anthropology of art and the anthropology of technology: a philistine, symmetric anthropology should be doubly indifferent to whether a recommender system “works,” since for a recommender system to work is both a technical and an aesthetic achievement. An anthropology of recommender systems, then, should not be concerned with determining how well they work, but rather with how they *come* to work:

> To create a new technology is to create not only a new artefact, but also a new world of social relations and myths in which definitions of what 'works' and is 'successful' are constructed by the same political relations the technology engenders. It could be objected, to be sure, that a technology either 'works' or it doesn't, but this objection obscures the mounting evidence that creating a 'successful' technology also requires creating and disseminating the very norms that define it as successful. (Pfaffenberger 1988, 249–250)

This is not so much a rejection of the idea that technologies "work" as it is a drawing into question of what it means to "work," the kind of critical dissolution advocated by Gell. This analytical goal is at odds with the common critical tendency to assume that recommender systems are made by culturally incompetent engineers and that their failings could be fixed if they were better informed. These casual critiques—why don't they use information about my friends? why don't they use information about the artists' social networks?—neglect the work that goes into making these systems work, the dense forest of negotiations through which something as apparently simple as "people like music because of how it sounds" has to pass in order to be implemented and the intricate response patterns that must be interpreted to decide if it "worked." They ignore the fact well-known by recommender systems engineers that "this works" is often shorthand for "this works for me," and the real challenge is to make a system that works for many.

Resonant Anthropology

In thinking about my relationship to the question of whether and how these systems "work," I had stumbled into a problem that has vexed the anthropology of expertise for some time: How should anthropological knowledge relate to the knowledge of its expert interlocutors? "On what basis," Dominic Boyer writes, "does the representative of one culture of expertise (the anthropologist) claim legitimate analytical jurisdiction over the members of another culture of expertise and how is this claim enacted?" (2008, 41). This topic has recently concerned anthropologists who have theorized alternative relationships among knowledge systems than explicans and explicandum. Anti-hierarchical, these orientations attempt to move horizontally, allowing for the coexistence of many competing knowledges and for the production of anthropological knowledge that does not necessarily displace other forms of knowledge: "lateral reason" (Maurer 2005), "collateral knowledge" (Riles 2011), working "athwart theory" (Helmreich 2009), and the "para-ethnographic" (Holmes and Marcus 2005) all bear this sideways orientation toward the work of explanation as well as in their interpretation of how knowledges interact elsewhere in the world.

This work is typically concerned with how the would-be objects of anthropological analysis have already produced semi-anthropological accounts of themselves or accounts which seem to compete with anthropological ones. The novelty here is not that people already account for themselves before the anthropologist arrives at the scene—Geertz described anthropological knowledge as "interpretations of interpretations," Boas described the materials anthropologists work with "secondary explanations" (Stocking 1982, 22), and interlocutors may always disagree with the readings provided by anthropologists. Rather, the difference is felt most acutely when the anthropologist comes "back home" and finds that the pre-existing analysis (what she may have wished to figure as an "emic" point of view) has "etic" credibility. I consider the tendency to presume that all knowledge about recommender systems is aimed toward making them work better to be an example of this kind of emic/etic blurring.

Gell's and Pfaffenberger's calls to ignore the question of whether art or technology "works" can read like calls for objectivity—for anthropologists to dispense with the mystifications that only seem to affect their subjects and not their selves. In another reading, though, methodological philistinism is not a claim to the view from nowhere, but rather a view from somewhere else. It is a means for producing anthropology's signature analytic effect: the bringing together of disparate worldviews in ways that destabilize the taken-for-granted in both of them. It is an attempt to not take the same things for granted that our interlocutors do, which inevitably means taking something else for granted.

My response to this situation was to develop what I've come to think of as a *resonant anthropology*. Resonance, I propose, offers a way to think about and connect the relationship between ethnographic objects and theories, anthropological knowledge and the knowledge of our interlocutors, and the rough-and-ready, everyday sense-making that obtains in the field. Such an approach is open to the embodied, affective, and environmental understanding that typifies sounded anthropology, but it is also open to alternatives that emerge unexpectedly from the noise of ethnographic fieldwork.

Scholars have turned to resonance—the sympathetic vibration of similar objects—when confronted with phenomena that are difficult to describe, or which seem to resist common, visually-informed understandings of understanding. Thus, "resonance" is frequently used to describe affective connections (Wikan 1992; Paerregaard 2002; Paasonen 2011) or, in the case of sociology, to describe how framings of cultural action intuitively "make sense" to enculturated minds (McDonnell 2014; see also Mary Douglas on self-evidence and gut reactions, Douglas 1972). In post-processualist archaeology, Ian Hodder has used "resonance" to describe a "process by which at a non-discursive level coherence occurs across domains" (Hodder 2012, 126), such that certain sets of objects and ideas seem to "go together"—he invokes Bourdieu's *Distinction* as an example of how a set of art objects come to be seen as linked. Resonance in these cases is a metaphor that captures something about lively, vibratory feelings of connection (Trower 2011) that persist outside of modernist logics of rationality.

However, resonance also offers the opportunity to think outside the dichotomy of an embodied, affective, immersive auditory knowledge and a transcendent, discursive, anatomizing visual knowledge. As Veit Erlmann recounts in his "history of modern aurality," *Reason and Resonance* (2010), many figures associated with the rise of enlightenment rationality dabbled with understandings of thought and logic rooted in ideas about resonance. Diderot, for example, described in 1769 the process of thought by analogy with vibrating strings:

> The sensitive vibrating string oscillates and resonates a long time after one has plucked it. It's this oscillation, this sort of inevitable resonance, that holds the present object, while our understanding is busy with the quality which is appropriate to it. But vibrating strings have yet another property—to make other strings quiver. And thus, a first idea recalls a second, and these two a third, then all three a fourth, and so it goes, without our being able to set a limit to the ideas

> that are aroused and linked in a philosopher who meditates or who listens to himself in silence and darkness. (quoted in Erlmann 2010, 9)

This account of thought as the progression of loose associations (rather than as a sequence of necessary entailments) seems remote from the kind of logical compulsions associated with enlightenment rationality or digital computation. Against the tendency to cede domains associated with rationality, logic, and mathematization and use the language of sound to define an oppositional space concerned with bodies, affect and environment, the persistence of these out-of-place understandings described by Erlmann and encountered in fieldwork provide an opportunity to stay with the mathematizers and produce new, resonant accounts of how they make sense. These accounts would not have to buy in to universalizing rationalist worldviews, nor would they be required to occupy theoretical ground defined in the negative against powerful formalizing knowledge practices. A resonant account would emphasize "adjacency, sympathy, and the collapse of the boundary between perceiver and perceived," as Erlmann suggests (2010, 10). According to the resonant theory of human hearing, perception is not the more-or-less-adequate mirroring of an exterior world, but rather a function of contiguity with it—the hearing organs set into sympathetic vibration with that which they hear. Theories resonate with the world they aim to describe, and they resonate with each other as they act and move in that world.

Erlmann's account of resonance resonates with the claim that Donna Haraway makes in "Situated Knowledges": "Feminist accountability requires a knowledge tuned to resonance, not to dichotomy" (Haraway 1988, 194–5). Whether we want it or not, "tones of extreme localization, of the intimately personal and individualized body, vibrate in the same field with global high tension emissions" (195). The critic cannot claim detachment, because she is always caught up in the global reach of technoscientific projects, making sense not from the outside, but from a position that is already partially shaped by the would-be objects of study. Resonance for Haraway links the apparently disparate "local" and "global," the material and the semiotic, the bodily and the scientific, and it provides a way to think about how anthropologists experience rapport and "complicity" (Marcus 1997) in the scene of fieldwork.

My understanding of "resonance" is informed by Haraway's argument, which necessitates discussing not only the objects of study but the disciplinary apparatus that has tuned into them. This understanding resonates with the "transductive ethnography" proposed by Stefan Helmreich in *Alien Ocean* and his insistence on working "athwart theory," through "an empirical itinerary of associations and relations" (Helmreich 2009, 23). To work athwart theory is to tack back and forth across the empirical and the conceptual, both as they are imagined by anthropologists and by their interlocutors; to approach ethnography transductively is to map out "the actual course that invention follows, which is neither inductive nor deductive but rather transductive, meaning that it corresponds to a discovery of the dimensions according to which a problematic can be defined," as Helmreich quotes Gilbert Simondon (Helmreich 2009, 243). Tying this discussion back to classic anthropological concerns, a resonant approach as I understand it does not take ethnographic experience as particularity to be inductively expanded toward the universal, nor does it take theory as universality to be deductively applied down to the particular.[15] Rather, it stays in the middle, among things, moving from the particular to the particular (Agamben 2009), or rather from object to object across apparent scales and domains, recognizing that the universal and the particular are not distinct planes of existence requiring induction or deduction to traverse, but rather are tangled up in resonant things that can work as theory or example, situationally.[16] In this, it resonates with other anthropological work that questions the taken-for-grantedness of particular/general, local/global scale and the modes of explanation such scales enable (e.g. Choy 2011; Zhan 2009; Tsing 2004).

I want to use resonance as a mode of tentative engagement with the theories of my interlocutors about how taste and technology work and relate to each other, to learn to hear as they do and in the process to learn something about learning, hearing, taste and technology. In the process, I have often found surprisingly resonant similarities between their work and the history of anthropological thought or between the design of recommender algorithms and the practice of ethnography. Resonance provides a way to

---

[15] In this I'm diverging slightly from another account of "resonance," which literary critic Northrop Frye gives in his book on interpreting the Bible: "Through resonance, a particular statement in a particular context acquires a universal significance" (Frye 1982, 217). Thus, stories and phrases from the Bible have the metaphorical capacity to take on extensive meanings beyond themselves. For Frye, resonance is another word for metaphor, "an identity of various things, not the sham uniformity in which all details are alike" (1982, 218). Identity without uniformity, similitude, is closer to what I am trying to mean by "resonance." But where Frye takes the power of resonance/metaphor to be its ability to expand particulars into universals, an anthropological approach would have to recognize the sociocultural contexts and power structures that grant specific particulars apparent universalizability and not others—why is it, for example, that phrases from the Bible are interpreted as though they could potentially apply to any situation, and who is licensed to do such interpreting?

[16] This approach thus resembles what Celia Lury, in the context of media sociology has described as an "amphibious" method, which moves across media, theory, and empirics,

understand such similarities between things that are not necessarily linked historically or causally (though sometimes they are); it offers a way to think about an anthropology premised on similitude rather than difference (Boellstorff 2005), in which the animating principle of the anthropology of science and technology is not that ethnographers and the ethnographized are bridging some ineluctable distance, but rather that near similarity and interaction (or resonance) is worth exploring as a motivator for ethnographic research.

It is in this spirit that I try to read the texts of my interlocutors from the field alongside the disciplinary texts of anthropology and its neighbors, not to use the discipline to explain the field, but rather to try reading them *as* each other—to not claim the role of cultural expert and evaluator for myself, but rather to see how my interlocutors manifest and understand cultural expertise. In what follows, I draw together resonant objects from my fieldwork experiences with scientists and engineers, the history of acoustic science, the history of social theory, and the practice of ethnography. Thus these chapters bring together the ethnological study of animal traps and the behaviorist imaginaries of developers in Silicon Valley, agriculture and the tending of unruly datasets by "data gardeners," a 19th-century theory of hearing and a computer that has maybe learned to identify heavy metal subgenres. The effect is intended to be something like what Lawrence Cohen calls "juxtapositional ethnography" in *No Aging in India,* drawing together an unruly set of theories, objects, theories taken as objects, and vice versa, "pushing the reader to make certain connections" (Cohen 1998, 8); or, the

---

floating in the middle of sociology's classic concerns with the micro-scale of interaction and the macro-scale of social facts and forces (Lury 2012).

technique of “dialectical montage” that Diane Nelson borrows from filmmaking in *Reckoning*, “laying two unlikely things like two faces beside each other” so that viewers make their own sense in the “cut” between them, “flooding these openings with our own meaning” (Nelson 2009, 73, 74). It is evocative and polyphonous (Tyler 1986), playing in the interferences and overtones of various objects, the rapid beating of frequencies nearly in tune with each other, the composition of complex waveforms, the rare and common coincidences of certain patterns. It is not meant to encapsulate or finalize its objects, but rather to carry forward some of their vibrations so that they might resonate with readers. This dissertation is a set of bells to be rung together: listen for the beating pattern of frequencies falling in and out of phase with each other and for the sympathetic vibrations they draw from your own set of resonant conceptual objects.

CHAPTER 2:
CAPTIVATING ALGORITHMS

Hooked

On an overcast day in the Bay Area, I was eating sushi with Mike, a longtime engineer at a personalized radio company. I had been living in San Francisco, visiting the offices of music recommendation companies and meeting with engineers in bars and coffeeshops. For the past decade, Mike had been responsible for his company's recommendation algorithm—the knot of code that decides what song to play next. This decision, he told me, is pivotal to the business: pick the wrong song and your users will leave, pick the right song and they'll stick around.

Over the years, his company's algorithm had become more complicated. It used to simply take a seed artist and then play music by similar artists. Now the algorithm was more elaborate, bundling together dozens of sub-algorithms under a master algorithm, which combined their outputs. The sub-algorithms were tuned to different listening styles: maybe a listener is adventurous, maybe she is not; maybe a listener cares mostly about genre, or maybe he cares mostly about popularity. Like many of my interlocutors, Mike described this variation by referring to stereotype users: the teen girl who wants to hear her pop music on repeat, the erudite jazz listener who is more open to exploration. Depending on how a listener respond to the recommendations, skipping songs, giving them thumbs-up or thumbs-down ratings, the balance among sub-algorithms shifts. The master algorithm is designed to maximize one metric above all else: the time users spend listening.

As listeners listen and the system collects more data, the logic of that sub-algorithm balancing changes: “Depending on where you are in your lifetime of interaction and experience with us,” Mike told me, “you get very different music experiences.” For long-term users with lots of listening data, the system can provide minutely personalized music choices that take into account long listening histories, pushing into more obscure musical territory with confidence. “But,” Mike continued, “if you’re in your first week of listening to us, we’re like, ‘Fuck that, play the hits.’ Play the shit you know they’re going to love to get them coming back. Get them addicted.”

“In the beginning, I’m just trying to get you hooked.”

Hooked, it turned out, was also the title of a self-published book by Silicon Valley entrepreneur and blogger Nir Eyal. Subtitled *How to Build Habit-Forming Products*, the book draws on behavioral economics and cognitive psychology to teach software startups how to get users to crave their products—to instinctually check their apps as a matter of habit rather than conscious choice. Successful products, according to Eyal, beat their competitors by making themselves “first-to-mind”—users “feel a pang of loneliness and before rational thought occurs, they are scrolling through their Facebook feeds.” The purpose of his book, Eyal writes, is to teach “not only what compels users to click, but what makes them tick” (Eyal 2013, Introduction).

Behavioral economics is ascendant in the startup world as companies like Mike’s seek to understand their users for the sake of “hooking” them. For many of the engineers I met in the field — people who would have once been criticized for overly rational assumptions about human activity—behavioral economics provided a frame in which to interpret human idiosyncrasies. Dan Ariely, the author of the popular book *Predictably Irrational* and former MIT professor, holds an annual summit in Silicon Valley called "Startuponomics," which pitches the principles of behavioral economics as strategies for companies looking to optimize their "product funnels," acquire users, or grow the enthusiasm of their employees. Like Mike's thinking about his listeners as entities to be "hooked," these uses of behavioral economics are all focused on forms of capture.

This chapter examines the relationship between developers' imaginations of users and the technical systems they design and build. I suggest that the goals evinced by Mike and *Hooked* — to keep users around, above all else—have come to define the purpose of music recommendation, as they have come to dominate the imagination of the software industry more broadly. Consequently, users are understood as entities to be captured, and recommenders can be understood as a kind of trap. Drawing on the anthropology of animal traps and its elaboration into a more general theory of captivation, we can see how understandings of prey—which include, but also exceed, behaviorist imaginaries—come to influence the technical form of traps. Captivation offers one frame for understanding how cultural and technical concerns mingle in the ongoing development of sociotechnical systems. To demonstrate this mingling, I describe a dominant understanding of users that circulated through the field during my research: music listeners vary in terms of their avidity for music. Recommender systems are thus tuned to avidity and are increasingly designed to capture the attention of the least avid listeners. Thinking of recommender systems as traps usefully draws into attention a set of concerns regarding the relationship between developers and users: the way the latter are understood by the former, the way these understandings are integrated into technical systems, and, ultimately, the ethics of captivation.

From Prediction to Retention

Recommender systems have not always been understood as technologies for hooking users. At their origin in the mid-1990s, recommender systems were effectively ratings predictors. To make recommendations, one simply had to predict what ratings a user would give a set of items, collect up the highest predicted ratings, and display them. Evaluating such a system meant comparing the predicted ratings to the eventual, actual ratings from users. Over the history of these systems, many competing techniques have emerged for making these comparisons, varying at practically every step of the process, from how ratings are collected, to how errors are calculated, to how those errors are aggregated together, to how the resulting recommendations are presented to users. A naive approach, for example, might average all the errors together: if a system predicted that a user would rate items 1, 3, and 5 stars, but they actually rated them 2, 4, and 1, then on average, it was off by 2 stars. Averaging those per-user errors together across the whole userbase produces a single number that represents accuracy and can be compared across different systems. However, such a metric treats all errors equally. One might care more about the accuracy of ratings at the extremes, where opinions are strong. On a five-point scale, predicting 4 instead of 5 stars might be considered more wrong than predicting 4 instead of 3. Or, one might penalize large errors exponentially more than small ones, since those can flip a rating from the “good” to “bad” side of the spectrum. There are many other such decisions to make and countless permutations of those decisions and their implementations (see Herlocker 2004 for a technical survey of evaluation techniques).

By the time I began my research in 2011, a standard paradigm for evaluating recommender systems had emerged, focusing on accurate ratings prediction and using one default error metric: root mean squared error, or RMSE. Like other metrics, RMSE aggregated the errors across all user-item ratings pairs into a single number, which effectively represented the difference between two ratings matrices. To calculate RMSE, first take the difference between a predicted and actual rating and square it. This amplifies larger errors (an error of 2 is now 4, while an error of 1 is still 1), and it makes all the numbers positive, ensuring that errors in opposite directions do not mathematically cancel each other out (otherwise being off by -5 on one rating and 5 on another might be rendered as an average error of 0). These squared values are all added together (taken from every user-item pair that ended up with an actual rating) and divided by the total number of user-item pairs (to get the average). Take the square root of the resulting number to return it to the scale of the original ratings, and you have RMSE: the root of the mean of the squared errors.

If you learned how to build a simple collaborative filter in an undergraduate computer science class, like the one I attended at UC Irvine in the spring of 2012, you would evaluate it with RMSE. If you presented a new collaborative filtering algorithm at the ACM RecSys conference, like the one I attended in Dublin later that year, you would likely evaluate it using RMSE. If you had developed a new method for evaluating recommendations, you would be obligated to compare it to RMSE. Root mean squared error was an integral part of the “normal science” (Kuhn 1962) of recommender systems that had coalesced by the late 2000s, taking as its main goal the improvement of predictive accuracy.

This research paradigm was epitomized and cemented by the Netflix Prize—a contest run by the online movie rental company from 2006 to 2009, offering $1M for an algorithm that could beat their Cinematch algorithm’s RMSE by 10%.[17] The contest was run “offline,” using a large set of ratings data Netflix had taken from its service. They provided competitors with a random sample of ratings from that data to train their systems, holding back a “test” set, on which the algorithms would be evaluated.[18] Cinematch could predict the held-out ratings with a 0.9514 RMSE. By the end of the contest in 2009, the winning algorithm, from the team “BellKor's Pragmatic Chaos,” had achieved 0.8567 RMSE and established the state of the art for predictive accuracy in collaborative filters.[19] The winning team was an ensemble of teams whose systems had performed well in earlier years of the contest. They combined their techniques into an “ensemble model,” which was essentially a weighted average of the various algorithms’ outputs. After the Netflix Prize, it was well-established common sense among recommender systems researchers that the best results were achieved through these systems that blended together the outputs of various techniques, rather than through

[17] Since RMSE is a measure of error, lower values are better: an RMSE of 0 would indicate that all ratings were predicted perfectly.
[18] This train/test split is standard practice in the development of machine learning systems to protect against “overfitting,” which is when a system performs very well on the data it was originally trained on, but fails on new data, indicating that its success is not based on generalizable factors, but on features specific to the training data.
[19] The worst RMSE was 3.1679, achieved on purpose (to be so bad is challenging) by the team Lanterne Rouge: http://www.netflixprize.com/community/viewtopic.php?id=336

the discovery of an ideal single technique. Ensemble models were pragmatically, rather than theoretically, driven: different sub-algorithms might be premised on distinct or even contradictory theories about the underlying rating patterns, but combining them produced the best performance.

RMSE abstracted the recommender problem into a matter of optimizing a single number, and this move endorsed a particular understanding of users (and cultural patterns more generally) as a set of numerical values to be predicted. The paradigm established by the Netflix Prize has thus become something of a punching bag for outside critics who see algorithms more generally as a rationalizing force that violently reduces culture to numbers. These systems represent "how central the accuracy of the recommendation system is to such organisations" (Beer 2013, 64) or "the enfolding of human thought, conduct, organization and expression into the logic of big data and large scale computation" (Striphas 2015, 396).[20] However, this paradigm was already faltering by the time of my fieldwork, encountering a crisis in how it understood users that would result a search for new evaluation paradigms that went "beyond RMSE," as the title of a 2012 workshop organized in part by Netflix's recommender researchers put it. These new evaluation schemes went hand-in-hand with a new way of understanding users.

[20] See also (Hallinan and Striphas 2014) for a review of the Netflix Prize and an emergent "algorithmic culture" aimed at a cultural studies audience. The Netflix Prize not only embodied a paradigm for recommender research, but for machine learning and data science more generally — the Australian company Kaggle would eventually offer a platform for competitions modeled on the Netflix Prize, allowing companies to run contests on their datasets.

As my interlocutors in the recommender systems research community told me, and as I saw in their conference presentations and conversations, the single-minded focus on RMSE had led to the neglect of actual users. Over time, steady improvements in RMSE began to taper off, trapped under what some researchers called a "glass ceiling." One graduate student explained to me at a recommender systems research conference that this was because people's preferences were "noisy" and unstable. In experiments where users rated items more than once, their own ratings often disagreed with each other, and those ratings could not be predicted any more precisely than they were held. An assumption that undergirded the use of RMSE and the development of collaborative filtering—that stable preferences were latent in the user, and could thus be predicted—was coming undone. For researchers drawn to recommender systems—who tended to focus on the algorithmic center of the system, rather than its interface or other features—new metrics proved further destabilizing: qualitative studies of users, which operationalized "satisfaction" through survey instruments rather than ratings accuracy, showed that improvements in RMSE did not inevitably lead to increased user satisfaction (Knijnenburg et al. 2012). Even worse, the effect of minor improvements in predictive accuracy was dwarfed by seemingly minute interface changes: one could achieve more dramatic improvements in user satisfaction by adding the phrase "Recommended for you" to the interface than by improving RMSE.

RMSE made the most sense for the offline evaluation of systems where users assigned explicit ratings: in these cases, researchers simply had to compare two matrices. But in commercial applications, recommendation had to be evaluated "online," as user ratings and activity happened, and this required a reorganization of computational labor. In domains like music, explicit ratings were less common, and recommenders had to rely on matrices of "implicit" ratings: did a user listen to a song all the way through or did they skip it? did they listen to it again later? These implicit signals introduced more uncertainty to the problem—a song skipped could mean many different things—and they were difficult to test offline: you would need a user's responses to recommended materials to know how well you were doing. Between the initial "training" phase and the ultimate "test," the recommender itself would intervene, causing measurable effects on people's behavior. These uncertainties highlighted the importance of interpretive work to determine what a signal meant.

In their report on the "Beyond RMSE" evaluation workshop, the organizers noted: "There seemed to be a general consensus on the inadequacy of RMSE as a proxy for user satisfaction, or an proper view on recommendation utility in general" (Amatrian et al. 2012, iv). When I attended RecSys again in 2014, in a Silicon Valley hotel overrun by industry representatives looking to recruit new talent, RMSE had fallen out of favor such that the head of recommendation at Netflix could present a slide with RMSE crossed out, banished to the evaluation of toy datasets in student projects, inapt for serious researchers or commercial application. He noted, as I had heard in many talks about the Netflix Prize over the course of my fieldwork, that the winning algorithm had never been implemented: it was far too inefficient, wringing out every last bit of predictive accuracy at the expense of computing time, and as Netflix moved from being a DVD rental company to a movie streaming service, they also moved away from a focus on predicting ratings. When users could easily stop a movie and pick another, giving them a solid estimated rating was less important than giving them an optimal set of movies from which they could pick. The goal of the recommender was no longer to provide the movie likely to be rated best, but rather to keep people watching movies on Netflix.com.

The key change in all of these shifts was the move to "user satisfaction" as the explicit and ultimate goal of recommender systems and a reconsideration of how satisfaction should be measured. This change was attended by a change in how users were imagined. Rather than presuming that satisfaction would follow from improvements to metrics like RMSE, researchers began to look for ways of measuring satisfaction that stemmed more directly from users. This move satisfied a number of demands at once: it addressed critics who argued that accuracy metrics like RMSE did not address user experience, it provided a way around the glass ceiling where RMSE had stalled, and it met the expectations of businesses who had little use for systems that were technically accurate but poorly received by the userbase. In a study of evaluative discourses among search engine developers, Elizabeth van Couvering (2007) found similar arguments, where a user-centered notion of "satisfaction" played a central role in the various technological schemas developers drew on to justify their work.[21]

Though some studies operationalized "satisfaction" as a qualitative measure, gathered from surveys, interviews, or focus groups, by and large the move to satisfaction meant a

[21] In the case of search engines, Van Couvering argues that other goals than satisfaction are possible and desirable: a search engine that aimed to serve the public good, for instance, might act differently with regard to controversial material than one tuned solely to satisfaction. Tarleton Gillespie (2012) makes a similar argument about "relevance algorithms" more generally—these systems are often used by publics to know

move from ratings and RMSE to the analysis of activity logs. Looking for satisfaction in the logs, the developers of recommender systems operationalized it as "retention." Or, perhaps it is the other way around: conveniently and not coincidentally, retention is not only a proxy for satisfaction—it is also an important metric for the venture capitalists who fund many of these companies: the astronomical growth on which venture capitalists stake their bets is a growth in userbase above all else. Across online media platforms, metrics like "dwell time" (how long a user spends on a service), monthly and daily active users (how many individuals have signed in during a given month or day), and the like have come to represent the "health" of a company to its investors. Where other metrics persist—RMSE and other accuracy metrics are still around, but in a supporting role—they are used in the quest for retention. As Mike explained to me how he wanted to "hook" users, he noted: "We're trying to really make your musical life good, so in the long term you'll come back."

The conflation of satisfaction and retention helped mediate the goals of developers, who were usually motivated by a strong desire to help listeners, and businesspeople, who wanted to acquire users. By calling this out as a conflation, I do not intend to suggest that it is wrong—it seems reasonable to suppose that a listener who listens to a service for an hour is more satisfied with it than someone who listens for only a minute. But nonetheless, this operationalization of satisfaction is a choice, driven by particular theories about listeners, investor imperatives, and an emergent common sense about how contemporary media companies should behave. Other end-goal metrics are

---

themselves, and being biased toward certain metrics like satisfaction can result in a skewed image.

possible and extant within the contemporary landscape of recommender systems, and they would also be choices, not necessarily good, right, or best: a recommender system could optimize for diversity or novelty at the expense of satisfaction, or it could operationalize satisfaction as something other than retention. But—and this is where the developers and critics of recommender systems often find themselves in agreement—appeals to user satisfaction have a moral power, and as a result, satisfaction is mobilized to justify a wide range of technical practices. Thus, Mike could, without any irony, argue that he was both working for the listener's best interest, while simultaneously trying to get them addicted.

The Anthropology of Traps

Critical scholars of the contemporary internet have suggested that we lack metaphors with which to apprehend the workings of algorithms (e.g. Crawford 2014). Given the central role they play in "hooking" users and the shifting of success metrics from prediction to retention, I am going to argue that it is useful to think of algorithmic recommender systems as a kind of *trap*. The figure of the trap brings into relief a set of concerns about recommendation that occupy critics and developers alike: the relationship between developers and users, the relationship between cultural understandings and technical functions, and the ethics of captivation. This move is not only motivated by the steady references to retention and capture I encountered in the field, but also by the world of anthropological theorizing that it brings into play.

Traps are a classic topic in the anthropology of technology. They populate the shelves of museums and the pages of ethnological reports—legibly technological bits of ingenuity and bricolage, assembled from string, sharpened sticks, bark, stones, and pits in the ground.[22] Early 20th century anthropological journals are dotted with brief descriptions of traps observed and collected around the world: conical Welsh eel-traps made of sticks and baited with worms (Peate 1934), a bamboo rat trap from the Nicobar Islands (Mookerji 1939), or pits dug along the Missouri river and baited with rabbits to trap birds of prey (Hrdlička 1916).[23] In other words, they seem quite unlike algorithms, which are supposed to be essentially immaterial abstract procedures, independent from any given implementation in code and certainly from any coarse materiality of the spearing, snaring, or smashing sort.

Through the history of anthropology, traps have offered empirical and metaphorical resources for theorists. In the early days of the discipline, evolutionist and diffusionist ethnologists took the materialities and mechanisms of traps as evidence for progress, migration, and invention. Henry Balfour, the first curator of Oxford's Pitt Rivers ethnological museum, took an interest in the distribution of Melanesian fish traps—long narrow baskets, lined with the thorny ribs of calamus plants—as evidence for a particular pattern of migration across the islands (Balfour 1925), only to receive word from E.E. Evans-Pritchard that similar traps were being used to catch rats in South

[22] Pits are hard for museums to accession, but so are original traps, which tend to unmake themselves in the act of trapping or to biodegrade and fade back into the environments where they are hidden. As a result, the traps registered in collections and photos are usually reproductions.
[23] I would be remiss if I did not describe this particular trap in more detail. According to Hrdlička, these pits not only contained rabbit bait, but also a man hidden underneath the bait, such that when the bird of prey swooped down, the man "would quickly seize the bird by the feet, pull it under the rafters into the dark, and wring its neck" (1916: 547).

Sudan (Balfour 1932), suggesting that the basic mechanism had likely been invented multiple times in multiple places.

Traps, like the ocean, blood, or trees, offer a large and pluripotent metaphorical vocabulary, which interpretive and symbolic anthropologists took up even as they turned away from the interest in materiality and mechanism which had caught earlier scholars' interest. So, we find Clifford Geertz describing culture as a web of significance in which people are suspended (1973), Roy Wagner recounting how signifying subjects find themselves "caught in Indra's net" (2001), or Pierre Smith analyzing rituals as "mind traps" (Smith 1984; see Halloy 2015). From these cases, we can get a sense of how culture *per se* is intimately tied to processes of capture, even as understandings of what culture *is* vary. A quick search through recent anthropological writing turns up studies of "hospitality traps" (Krögel 2010), "credit traps" (Williams 2004), and the curious case of "captive guests" among the Nuosu people of southwest China, who conceive a webbed cosmology in which individuals possess (or are possessed by) "soul spiders" that attempt to capture each other (Swancutt 2012). These latter references revive, in various guises, the modes of reciprocal obligation and entanglement[24] that anthropologists have long studied in the wake of *The Gift* (Mauss 1990).[25] People, their thoughts and actions, are

[24] Rey Chow, in her *Entanglements, or Transmedial Thinking about Capture*, argues that the trap is "an archetypal epistemic or representational device, a dispositif (in Foucault's terms), perhaps" (Chow 2012, 45). Her concern, in dialogue with Jacques Ranciere and Alfred Gell (to whom I turn in a moment), is the boundary between art and non-art, on which the figure of the trap seems to operate.

[25] Perhaps these gossamer cultural ontologies are relatives of Tim Ingold's polemical social theory "for arthropods" called SPIDER, meant to emphasize emplaced relationality over the discreteness of actants. Where actor-network theory is for ANTs, SPIDERs know that **S**killed **P**ractice **I**nvolves **D**evelopmentally **E**mbodied

caught up in culture, while they try to capture each other into relationships and ongoing projects.

But if traps have predominantly worked in the discipline as either artifacts or metaphors—in the study of technology as a matter of mechanism or sociality as a matter of obligation—we occasionally find them as material-semiotic figurations, plying together the technical and the symbolic (Haraway 1994). In a striking article from the 1900 volume of *American Anthropologist*, Smithsonian curator Otis Mason catalogued the variety of traps used by indigenous people across North America. These traps, he wrote, were "ingenious mechanical combinations" (Mason 1900, 659), exceedingly complex in comparison to ordinary tools: "automatic action of the most delicate sort is seen in the traps themselves, involving the harnessing of some natural force, current, weight, spring, and so on, to do man's work" (Mason 1900, 658). But the complexity of these traps was not limited to their physical form—it extended to the "psychology" of the animals to be trapped:

> the trap itself is an invention in which are embodied most careful studies in animal mentation and habits—the hunter must know for each species its food, its likes and dislikes, its weaknesses and foibles. A trap in this connection is an ambuscade, a deceit, a temptation, an irresistible allurement: it is strategy. (Mason 1900, 659)

In the trap, according to Mason, the psychology of the trapper and the animal met in technical form—"the thought of the hunter had to be locked up in its parts ready to

---

Responsiveness (Ingold 2008, 215). I will leave the anthropology of spiders to less arachnophobic scholars.

spring into efficiency at a touch," anticipating the animal whose actions were a necessary part of the trap's functioning (Mason 1900, 660).

This understanding of traps as the material interaction of psychologies re-appeared a century later in a late essay by Alfred Gell: "Vogel's Net: Traps as Artworks and Artworks as Traps" (1996). Gell's interest in traps was stimulated by his interest in defining "art" cross-culturally, surmounting the distinction between artworks and mere artifacts. Prefiguring his semiotic theory of art that would be posthumously released as *Art and Agency* (Gell 1998), "Vogel's Net" advances the argument that what unites traps and art is their tangle of agencies—their apparent power to act as "a nexus of intentionalities between hunters and prey animals [or artists and audiences], via material forms and mechanisms" (Gell 1996, 29). A trap is "both a model of its creator, the hunter, and a model of its victim, the prey animal. But more than this, the trap embodies a scenario, which is the dramatic nexus that binds these two protagonists together, and which aligns them in time and space" (Gell 1996, 27).

Gell argues that artworks can be considered as traps, and traps as art, because both are technologies of *captivation*. It is helpful here to keep in mind the blurry, older sense of "captivation" referring to the capture of both mind and body. The Cartesian distinction, reinforced by positive and negative connotations, has split the definition of the word in two over time: it is perfectly acceptable for a musician to captivate her audience's minds, but not their bodies, for example. It is this generic sense of captivation—mental and physical—that finds a connection between the making of recommender algorithms and the making of traps. For Gell, this understanding links Duchampian urinals, the realism of the Dutch masters, the dazzling prows of Trobriand canoes, and thorn-ribbed Melanesian baskets, all of which, through varying means, captivate.

Daniel Miller applied Gell's theory of captivation to internet technologies in his article "The Fame of Trinis: Websites as Traps" (Miller 2000). In it, he argues that Trinidadian home pages function like traps, using aesthetic qualities to keep visitors around. Miller's article came out the same year that pop social theorist Malcolm Gladwell's breakout book *The Tipping Point* (2000) was published, containing a chapter titled "The Stickiness Factor," which described how successful media properties and trends could "stick" in audiences minds. This concept, though somewhat reversed as it stuck to its readers' minds, would become extraordinarily popular for talking about websites: companies sought to design their websites to be "sticky," so that visitors would stay around (and see more advertisements).[26] Recommender systems, with their focus on retention, can be seen as a further development in this vein.

The purpose of analogizing recommender systems to traps is not just to note a surprising resonance between the two. Rather, it is to redirect our attention, to suggest features to investigate and interpretations to attempt. The first of these is the way that traps, as Gell notes, "can be regarded as texts on animal behaviour" (Gell 1996, 27). He translates a story recounted by the anthropologist of religion Pascal Boyer, which was

---

[26] The concept was so popular that it spawned a backlash: in their book *Spreadable Media*, Henry Jenkins, Sam Ford, and Joshua Green argue that designing for stickiness can turn a website into "a virtual 'roach motel'" (Jenkins et al. 2013, 6; i.e. a trap), and suggested that instead media companies should structure their content to be easily shared, remixed, and spread. (The term "spreadability" was an attempt to relocate the

told to him by Ze, a Fang epic chanter in Cameroon, about the trapping practices of a neighboring group:

> In my youth I got to know the Pygmies well. The Pygmies belong to the forest, they are not village people like us [...] I often went hunting with the Pygmies, they have special traps for every kind of animal, that is why they obtain so much game. They have a special trap for chimpanzees, because chimpanzees are like human beings: when they have a problem, they stop and think about what to do, instead of just running off and crying out. You cannot catch a chimpanzee with a snare because he does not run away [and thus does not pull on the running-knot]. So the Pygmies have devised a special trap with a thread, which catches on the arm of the chimpanzee. The thread is very thin and the chimpanzee thinks it can get away. Instead of breaking the thread, it pulls on it very gently to see what will happen then. At that moment the bundle with the poisoned arrow falls down on it, because it has not run away like a stupid animal, like an antelope would. (Gell 1996: 25; Boyer 1988: 55–56)

Chimpanzees are curious creatures, while antelope are running creatures, and traps for each are designed accordingly. But there is more to it than that. For these traps do not simply reflect the behavior of animals, they reflect the human *understanding* of this behavior, which is often tied to higher order concerns than straightforward chimpanzee or antelope ethology. For example, for Ze and his fellow chanters of *mvet*, magical epics, traps and epics are alike, and the plight of the chimpanzee, who seeks knowledge at the expense of his life, illustrates "the basic Faustian problem about knowledge" (Gell 1996:

---

agency in this process from the circulating material—where agency is located in the term "virality"—to the people doing the spreading.)

25). As Boyer writes of wisdom, “you think you can get hold of it, but it escapes, and it is you who gets caught” (Gell 1996: 24–25). Traps are not only traps—they are “a master metaphor of very deep significance” (Gell 1996: 25). This is not only true for Ze in Cameroon or for anthropologists in search of metaphors, but for the interpreters and builders of other traps as well: the work of captivation refracts broader questions about knowledge, culture, and technology.

Traps and recommender algorithms are anticipatory technologies, which by necessity mingle technical considerations with theories about the entities anticipated. In the design and construction of a trap, it is possible to read theories about the specific entities in play—an idea about how chimpanzees pay attention, for example—but it is also possible to sense a theory about what animals are like more generally and how they relate to their trappers. Franck Cochoy, in an article on marketers, gives an account of “captation”—his term aimed at designating captivation in a generic sense. He argues that captation is uniquely well-suited to demonstrating the entanglement of sociotechnical assemblages (dispositifs) and socially patterned tendencies of action (dispositions): “the captation of publics consists in putting to work dispositifs which attempt to profit from dispositions that one attributes to persons in order to shift their trajectories” (Cochoy 2007, 204). In other words, when people attempt to capture others, they do so by anticipating particular tendencies and making adjustments to “ad hoc dispositifs” (Cochoy 2007, 207): “We could say that captation is about observing the path of a target, to anticipate its trajectory, to try and join up with it, to accompany it, to encircle it and to guess it in order to then attract it or intercept it” (Cochoy 2007, 212). These anticipated dispositions or trajectories are not necessarily based on academically legitimated social theory, and they do not need to be — the anticipatory structures into which they are built can perform their theories into existence, structuring the scenarios in which trapper, prey, theory, and technique meet in space and time. These "spontaneous sociologies" (Cochoy 2007, 211) can be as influential as they are diverse.

When critics (e.g. Beer 2013; Striphas 2015) worry about the shaping effects of algorithms on culture, they worry about the performative capacity of these ad hoc theories (MacKenzie 2006). Because traps do not reflect an image of their prey, but rather try to affect them (to "shift their trajectories" as Cochoy puts it), their accuracy is somewhat beside the point: regardless of whether they are correct, they intervene. The shift from predictive accuracy to user retention in the evaluation of recommender systems can be read as the recommender research community coming to realize this, placing more emphasis on the recommender as an interactive technology, unfolding in time. While algorithmic systems embody theories about how the world works (MacKenzie 2006), our understanding of those theories needs to extend beyond whether they are correct or not, since they can vary arbitrarily and the system will still "function" (this being determined within evaluation schemes that are themselves informed by those theories).

The effect of arbitrary cultural theories on the design of traps is evident in the work of the French anthropologist of technology Pierre Lemonnier (Lemonnier 2012). Lemonnier conducted fieldwork with the Ankave Anga in highland Papua New Guinea, who, like other groups in their region, catch freshwater eels using traps. These traps are made from a tube of rolled-up bark, tied together with rattan. The back end of the tube is blocked off, while the front is open, allowing an eel to swim in. A trigger at the back of the tube is connected by a long string and flexible stick to a trap door at the front. To set the trap, one ties a live frog to the trigger, bends the stick, and places it in the river. When an eel, tempted by the frog, swims into the trap and tries to eat, it trips the catch and the door snaps closed, sealing it inside.

The distinctive feature of Ankave traps is that they are heavily reinforced. Although similar traps are found among groups up and down the river, catching the same eels with no special difficulty, the Ankave traps are bound with extra rolls of bark and more rattan loops. This poses a puzzle for Lemonnier: Why would the Ankave make traps that are much stronger than they technically need to be, at the expense of time and materials? Lemonnier's answer lies in the cultural significance of eels to the Ankave: Although all sorts of people in Papua New Guinea eat eels, the Ankave have a "particular interest" in them (Lemonnier 2012, 56). The eating of eels plays a central role in their mortuary rituals, and they are considered incredibly potent. This potency stems from an origin myth that links eels to men's penises and to the bond between people and the land on which they live. The rituals that attend the making of traps reinforce this symbolism, through scripted roles played by men and women as builders, testers, and guarantors of traps' efficacy. Eel traps are thus not "merely" bits of technical ingenuity, but are rather densely signifying objects in an overarching system of meaning that links together the power of the ancestors, the connection of men to the land, and the management of the dead.

Lemonnier argues that archaeological common sense takes “cultural” influence on technologies to manifest only in their incidental stylistic features—the decorative leaves tucked into the rattan binding, arbitrary with regard to the trap’s functional elements. Drawing on Ankave material, Lemonnier’s striking conclusion is that cultural arbitrariness can appear even in a technology’s “technical” aspects, like its strength. The functioning of the trap is not determined entirely by its adequacy to the “actual” behavior of the entity to be trapped, which might be gathered only with the right science to figure it out—many possible arrangements of traps will find their basic felicity conditions met, and they may even have a part in shaping behaviors to meet them. As Sandra Harding noted regarding the history of science: “nature says ‘yea’ to many competing and, from our perspective, quite fantastic accounts of its regularities and their underlying causal tendencies; our best theories are only consistent with nature, not uniquely coherent with natural laws that are ‘out there’ for our detection” (Harding 1995, 346).

One may be inclined to agree with the Fang chanter Ze that traps to catch chimpanzee and antelope are different because chimpanzee and antelope are different. But these different traps that catch the same eels pose the question of cultural influence more explicitly. If we tried to read the “text” on animal behavior constituted by the trap, as Gell suggests, we would be led astray without the context from which it emerged: the eels near the Ankave are no different than the eels downstream, save for the cultural worlds they swim through. This blending of “technical” and “cultural” concerns supports Bryan Pfaffenberger’s suggestion outlined in the last chapter that anthropologists remain agnostic about whether a technical practice “works” or not, for the sake of being able to see the broader contours of the sociotechnical system. An explanation that attempted to isolate functional concerns from cultural ones could not account for this difference. In other words, to understand how traps come to work as they do, requires more than understanding than their basic mechanisms—the snapping of levers or tightening of strings—it requires understanding their makers' cultural worlds.

This suggests a new way to pay attention to the making of algorithmic recommender systems. They are not merely a matter of optimizing a metric or implementing a technical fix—they are attempts to *captivate* people. Like other forms of captivation, they are shaped by cultural theories about the entities they seek to captivate (whose trajectories they seek to bend, in Cochoy's terms). These theories are "cultural" both in the vernacular sense that they concern music and taste, but also in the sense that they can vary arbitrarily: if there were another village up the river from Silicon Valley, one might find other theories guiding the construction of traps. In the balance of this chapter, I outline in detail the dominant theory for understanding listeners I encountered during my fieldwork: that they vary primarily in terms of their avidity for music.

Avidity

For the Fang chanter Ze and the trappers he described, animals varied in terms of their intelligence: chimpanzees were smart, antelope stupid, and traps to catch them had to vary accordingly. This is not merely the truism that technology is built with models of the world in mind and that its efficacy is directly related to the adequacy of these models. Rather, for Ze, traps became an occasion for thinking about the nature of knowledge, intelligence, and wisdom: they require knowledge to make, they interpellate their prey with regard to their intelligence, and the whole enterprise of trapping is a materialized metaphor for the acquisition and disposition of wisdom.

For the developers of music recommender systems, we can discern a similar organizing concept in their concern for the intensity with which people pursue, enjoy, or learn about music. I term this *avidity*, to capture a set of related ideas about enthusiasm, appreciation, and knowledge. In the field, this idea went by a variety of names: listeners might be "high intent" or "low intent"; they might be high or low "engagement"; or, colloquially, they might be more or less "into music." Avidity is central to how developers understand users (and their variability); it organizes developers' understanding of the difference between themselves and their users; and it undergirds the purposes of recommendation more generally.

I first learned about the importance of avidity in one of the earliest interviews I conducted for this project in 2011, with Peter, a senior engineer at a company I'll call Whisper. Whisper is what media scholar Jeremy Morris calls an "infomediary"—it provides music data infrastructure to other companies (Morris 2015). While those companies work on things like acquiring licenses and serving audio streams, Whisper focuses on the information flows that are the guts of the recommender system—listening data, music metadata, other information scraped from the web, playlisting algorithms, and so on, all made available through an API.

I would come to know Peter much better over the next few years: we attended the same conferences, we had follow-up interviews, and eventually, I would spend several months as an intern at Whisper, working under him. Peter was a well known figure in the world of music recommendation, in both industry and academia. He spoke regularly at conferences and hackathons, and wrote widely read posts on his blog.

When we first met, in Whisper's office, I asked him what he thought were the biggest challenges facing the developers of music recommenders. He told me that one of the most challenging problems was that listeners varied, and different types of listeners wanted different things out of a recommender system. He recalled a study he had read about that defined "a pyramid of listeners." At the bottom of the pyramid were the "musically indifferent"—people who did not really care about music—who "wouldn't care if music went away." He estimated that these musically indifferent people made up about 40% of the population. Above them in the pyramid—in decreasing size and increasing avidity—were "casuals," "enthusiasts," and then "savants" at the very top. Indifferents were sometimes called "Starbucks listeners," because they would have been happy to just get their music from the compilations sold on the counter at Starbucks coffee shops. Savants, however, were different. "Their whole identity is wrapped up in music," he told me. Later, he would say that they "live for music"—they "do the bar quizzes and all that stuff," avidly pursuing both new music and information about it.

When we talked in 2011, Peter couldn't remember where the pyramid had come from, but over the next few years, it kept popping up in my field notes: mentioned on stage at a music informatics conference in Porto or in a Skype call with a computer science graduate student in the UK. Eventually, I traced it back to a book: *Net, Blogs and Rock 'n' Roll* (2007), written by music industry consultant David Jennings. Jennings had taken the pyramid from a market research study conducted in 2003 and 2005 by UK media conglomerate EMAP called "Project Phoenix," which surveyed listening behavior in Britain. In his 2007 book, Jennings is circumspect about the study's generalizability, noting that it is a "snapshot" of a particular moment and place. But as we might expect (Latour and Woolgar 1979), these modalities were deleted as the pyramid traveled. As people like Peter used it to think with, reconciling it with their own intuitions about music listening, it was re-shaped into a generic truth about audiences.

Throughout my fieldwork, I saw how avidity absorbed and came to stand for many types of difference. At the bottom of the pyramid, we might find the stereotyped unserious fourteen-year-old girl who Mike referred to in the beginning of this chapter, who only wants to listen to what's popular, or the older listener in the midwest, who knows what she likes and wants nothing else. It also displaced the popular strategy of distinguishing listeners by their favorite genres: metal and electronica fans were known for their intense taxonomizing, while pop and country fans were more likely to be musically indifferent. Perhaps most significantly for companies like Whisper and their corporate customers, avidity was correlated with class and the willingness or ability to spend. Although savants were rare, they would spend significant amounts of money on music. But indifferents represented the largest market opportunity: "If you're looking for the big bucks," Peter remarked, you should try to capture the attention of the musically indifferent. This group had become the object of significant attention among the developers of music recommendation.

In the day-to-day work of engineers, avidity came to represent the distinction between company insiders and outsiders. Although the people who worked at companies like Whisper were predominantly white, English-speaking men between 25 and 40, the difference between insiders and outsiders—between engineers and users—was most often explained to me in terms of avidity. In interviews, engineers told me how they had cultivated interests in music performance, DJing, or fandom for long times; they were deeply knowledgeable about the minutiae of subgenres of heavy metal, they listened to obscure music from around the world (occasionally with formal ethnomusicological training), or they were generalists with, as one engineer put it, "a wikipedia-like knowledge" about music. The target market for recommender systems, however, consisted of people farther down the pyramid, definitionally less interested in or knowledgeable about music.[27] This posed a challenge that developers of recommender system were acutely aware of: as "savants," they did not relate to music in the same way that "indifferents" did. Engineers tended to be, as one Dublin infomediary named itself, "Music and Data Geeks."[28]

[27] This was not exclusively the case: often, engineers would build systems to fulfill certain desires that were thought to be signatures of more avid listeners: elaborate visualizations, tools for browsing obscure connections or surfacing rarely listened to music, and so on. These, however, were typically understood to be the kind of side projects that contemporary software engineers partake in as part of a more general programming habitus, at events like hackathons.
[28] One of my academic interlocutors joked that Peter was a "trans-savant," because although he was an extraordinarily avid music listener and creator of tools to aid browsing, his preferred genres—pop music, classic progressive rock, and pop punk—were more like those of less avid listeners. It was as though he had broken through the end of the scale and looped back around to the beginning.

On the last day of the 2014 RecSys conference, held in a surprisingly isolated Silicon Valley hotel in the thoroughly planned Foster City, I was drinking in the hotel bar with a few of my long-term fieldwork contacts. Some of them had recently moved from academia into industry, and they were arguing about the difference between themselves and the users of their services. Oliver, who now described himself as a “data plumber,” claimed that engineers have niche musical tastes—this made it hard for them to evaluate the experience for an average user, who would most likely want music from the more popular end of the catalog. "It’s hard to recommend shitty music to people who want shitty music,” he said, pointing to the contradiction between two evaluative schemes: his own idea of what was good, and the criteria for evaluating a recommender. Seth, a research scientist at a different company, disagreed: “We’re not tastemakers,” he said. The recommender should give people what they want, even if it offended the engineers’ sensibilities. That being said, he attributed the success of his relatively new employer to the savant status of the company culture: “everyone in the company is really fucking cool.” In public debates that had grown around Silicon Valley in the previous few years, “culture fit”—the idea that employees should not only have relevant skills but also fit more subjectively into a company “culture”—had become an object of critique for its homogenizing force: workplaces that hired people like the people already there tended to hire more young, white men. Seth mobilized the idea of company culture not in reference to this broader discourse, but with regard to how critical he saw it that his co-workers were cool: “you can’t risk polluting the culture.” The boundary between those who worked for the company and those who did not was understood in terms of this taste-based stratification.

The problem of insider-outsider distinctions is not new or unique to companies like Whisper. The field of human-computer interaction has its origins in the early 1980s in attempts to ameliorate the problem that a typical computer system was "unusable by anyone other than the people who built it" (Moggridge 2007).[29] In his canonical STS paper "Configuring the User" (1991), Steve Woolgar described how ideal "user" and "engineer" types served to reinforce the insider-outsider distinction for employees at a computer company in the late 1980s. As it did for Whisper, this distinction, coupled with the desire to keep work-in-progress secret, made it difficult for "insiders" to access the perspectives of "outsiders." Critics have described how such situations can result in a turn to the "I-methodology," in which designers and engineers imagine themselves as users (Oudshoorn et al. 2004), or to designing for their "imaginary friends"—sketchy personas that inadequately reflect actual users (Massanari 2010). These processes result in technical systems that neglect the situations of groups underrepresented in technical professions, and the contemporary tech industry is riddled with such products and features: "real name" policies that put victims of stalking or abuse at risk (boyd 2012), persistent social network profiles that stymie those making gender or other transitions (Haimson et al. 2015), facial recognition systems that cannot see non-white faces (Phillips 2011), speech recognition systems that cannot hear women's voices (Margolis and Fisher 2002), and so on.

Critics of such practices often suggest that they might be improved by improving information flows: through user research that provides more adequate knowledge about how technology fits into contexts of use, or which is more closely integrated into the

[29] Thanks to Ellie Harmon for bringing this reference to my attention.

design process. Woolgar describes how this approach was taken up by the user testing division at his field site: it was their job to rein in engineers' speculations (and derisive comments) about users, maintaining that proper knowledge about users could only come from proper research. This was a recognition of what channels actually existed for bringing outside knowledge in, but it was also a political strategy, asserting the power and purview of one company division over another. If the engineers wanted to avoid building for their imaginary friends or themselves, the alternative would be to turn to legitimated techniques for producing social knowledge.

For engineers at Whisper and companies like it, which have benefited from decades of work by HCI researchers to spread these critiques, the situation was more ambivalent. They were quite aware of the limitations of the I-methodology. Engineers frequently reminded me and each other that, as enthusiasts, they could not simply imagine themselves in users' shoes or pretend that their desires for features were widely shared. While I-methodology undoubtedly persisted in the ad hoc pragmatics of making software—in the countless "sniff tests" and informal evaluations of functionality that mark the working day (see the discussion of listening in Chapter 4)—it was regularly brought into focus as a problematic tendency to be resisted. Rather than imagining oneself as the user, one had to imagine the user as someone unlike the self. These imaginations might draw on user research at a distance (like the Project Phoenix pyramid or similar efforts undertaken by internal research teams), but they just as often drew on personal experience with friends and family, or on speculation through cultural stereotypes.

This question of how and whether engineers can take on the perspective of users resonates with Rane Willerslev's account of trapping by Siberian Yukaghirs. He quotes the hunter Taishin Arkadi, describing how trappers must learn to "think like" the animals they seek to trap: "The character of the sable is curiosity, pure curiosity. To place your traps well, you must be curious like the sable" (Willerslev 2007, 91). "Trapping," Willerslev explains, "involves a kind of mental projection by which the hunter seeks to place himself imaginatively within the character of the animal, matching that which is unique about it" (Willerslev 2007, 91). These mental projections are not necessarily accurate, and, more importantly, they do not have to be accurate to "work"—the coupling between user imaginaries and actual users is loose and accommodating, with enough room for arbitrariness that many configurations of imagination and evaluation will return results that work.

The practices Willerslev describes are congruent with the subjectivity-disavowing practices of the engineers I talked with about their work, who simultaneously drew on their own personal experience and musical expertise while trying to think outside of it: "perspectivism among Yukaghirs is not really about moving from one point of view to another. Rather, it is about not surrendering to a single point of view. It is concerned with action in between identities, in that double negative field" (Willerslev 2007, 110). Like Yukaghir "soul hunters," the developers of recommender systems locate themselves in the middle of a double negation: they need to think not like engineers while recognizing that they are also not like users.

For many of my interlocutors, this double negation was managed by references to science, which could provide a putative position outside of individual perspectives. On a cross-country video conference in summer 2013 with Tom, a product manager responsible for "audience understanding," he described his company's data collection efforts: "We don't interview users," he said. Instead, he told me, they used aggregated listening data, which reflected actual listening behavior, rather than claimed listening behavior.[30] "We think we have real science here." Science, in this case, referred to the kind of knowledge that could be derived from activity logs—not from direct engagement with users. Instead of turning to user research divisions to learn what their users were like, my interlocutors turned instead to the logs of listening data accumulating on their servers. Even for companies with traditional user research divisions, who conducted interviews, led focus groups, and occasionally pursued short-term ethnographic projects about listener behavior, these logs of activity data offered engineers a readily accessible way to understand users at large and small scale. The work of dedicated user researchers increasingly involved interpreting and reporting on this data, but the availability of this data to anyone in the company and the technical expertise required to parse it meant that engineers had become newly empowered to produce accounts of user tendencies themselves, reducing their reliance on a distinct company division, the importance of which receded accordingly. They enabled the detection of statistical patterns from all the users of a system while at the same time making it possible to recreate in some detail all the interactions of a single listening session: tracks played, skipped, and so on.

[30] Note the similarity between this argument for data mining and arguments for participant observation: what people say is not necessarily what they do. Tom would later advocate for a project comparing publicly stated "likes" on platforms like Facebook to actual listening behavior, which he took to be a listener's true likes.

Represented as aggregations of play events, the first striking feature of users was the variation in their quantities of listening: here, a set of users listening to the same album on repeat, there, someone who listened to a few songs and signed off, there someone who never listens to the same song twice. Knowing users through the logs and understanding them in terms of avidity went hand-in-hand: avid listeners had larger, more diverse taste profiles, while indifferents would scarcely register. It is perhaps not surprising that, when users were known primarily through the various bulk of their listening data, they were understood to vary by just how much data they generated. These logs of behavioral data should bring to mind the behaviorism of *Hooked*: while my interlocutors would argue that music listening was highly context-dependent, listeners were nonetheless referred to primarily with regard their propensity to click.

Tom referred to the listening behavior aggregated in the logs as a user's "musical identity." When Whisper introduced a feature that aggregated these events together, they were called "taste profiles." These names indicate how listeners, their taste, and their listening history had come to be seen as identical within the frame of recommender development. Although engineers would bemoan the fact that their own recommendations had been broken by listening histories full of music they had to listen to for feature testing, they presumed that the normal listening behavior of normal users reflected their "musical identity." Perhaps for ordinary users, the mapping of listening history to taste would be close enough, and irregularities would fade away over time as more enduring tendencies crowded them out. To be a listener in this paradigm was to be a collection of listening events. This is not to say that my interlocutors were unaware that people had lives beyond the service or listening activity that was not necessarily caught in the logs, but rather that, for the limited purpose of making the recommender work, this reduction of the listener to the log was considered pragmatic and reasonable.

In practice, the four-level Project Phoenix pyramid often collapsed to a binary distinction that was readily visible in the logs: avid listeners were “lean-forward,” fiddling with settings, actively engaging with the interface, browsing, and skipping songs, while less avid listeners were “lean-back,” looking to start the music and then leave it alone. This language was even more widespread than the details of the pyramid, encapsulating how people thought about entire services: an algorithmic radio service like Pandora was essentially lean-back, while the MP3 blog aggregator The Hype Machine was lean-forward, encouraging enthusiasts to browse new and obscure tracks posted to the internet. On-demand streaming services like Spotify or Rdio, with their large catalogs, boasted features aimed at both kinds of use: listeners could start algorithmic radio stations and lean back or lean forward and browse through the catalog, aided by algorithmically generated “related artists” and other classifications to guide navigation.

With the simplification of the lean-back/lean-forward binary came a more situational understanding of listeners. Where labels like “savant” and “indifferent” smacked of an unseemly essentialism, with moralizing overtones, “lean-back” and “lean-forward” were typically used to describe situations. The same person, depending on her context, might want different modes of interaction: while exercising, she would not want to interact, but while looking for new music at her desk, she might be lean-forward. As Tom put it:

"one listener is really many listeners." There was regular slippage between the situational and the essentialist understanding of listeners: when pressed, my interlocutors would acknowledge that leaning forward and back were contextually determined attitudes, but while going about their work, "lean-back" would occasionally lapse into the same essentializing sense as "indifferent," standing for a group's basic tendencies rather than the contingency of particular moments.

Zero UI

Peter argued that avidity manifested in a user's willingness to interact with a system. Savants or lean-forward listeners might make a hundred interactions in a listening session, hand-picking every song, while lean-back casual listeners might only want to interact once—to pick a "seed" artist for an algorithmic radio station, for instance. The problem was both that different kinds of listeners wanted different things—savants might want aids to browsing, while casual listeners only wanted to pick a single artist—and that different kinds of listeners provided dramatically different amounts of data about themselves. "In any of these four sectors [of the pyramid]," Peter told me, "it's a different ballgame in how you want to engage them." Like chimpanzees and antelope, variously avid listeners required special traps to be caught.

The biggest challenge, however, was posed by the indifferent listeners at the bottom of the pyramid. Although they represented the largest potential market, they were definitionally unwilling to interact with a system. This meant that they produced very little data. For a system reliant on interaction data, a user who doesn't want to interact is a mystery. The challenge of making recommendations with very little interactional data is known in the recommender systems research community as the “cold start problem”: When a user first signs on to a service, there is not enough data about them to correlate with anything, and thus to provide recommendations. (The same is true for new items to be recommended.) Recommender algorithms that work well for people with a longer history of interaction fail for these users who don’t yet have interaction histories. Thus, they are typically supplemented by a range of “bootstrapping” techniques. In the beginning of this chapter, when Mike suggested that a system should just “play the hits” for new users, he was referring to one such bootstrapping technique: popular music is more likely to be liked, and this was thought to be especially true for indifferents.

On his blog, Peter proposed another solution to this problem: if indifferents didn’t want to interact, maybe they didn’t have to. He called his idea “Zero UI,” a special kind of trap, “a recommender that can capture the attention of indifferent listeners.” Thanks to the proliferation of sensor-packed smartphones and an ecology of data-sharing applications, it was possible to collect a wide-ranging set of “implicit signals” about listeners, their contexts, and, perhaps, their taste. These signals ranged from location to the time of day to basic demographic information collected from social media profiles. “When listeners change the volume, when they skip songs, when they search or stop listening, they tell us about their taste.” Other potential signals included the weather, listening history from other services, or the events on one’s calendar. “My phone knows that I’m late for a meeting. Maybe it knows my the favorite songs of the people I’m meeting with.” From a motion-tracking fitness sensor, your phone may know if you’re running or riding in a car. From social media accounts, it may know your age, gender, and regional origin. Thanks to the emergence of big data infrastructures, all of these signals could be logged, aggregated, and correlated with each other, producing a data profile even of listeners who had not yet touched a button in the music player.

This passive interaction required no interface, other than perhaps a play button that would automatically play the “right” music. The Zero UI ideal existed as a kind of limit point for algorithmic recommendation, where everything about a music streaming service was reduced to nothing, except for the recommender at the center. It sought to overcome the limitations of what could be inferred from data in the logs not by changing the imagined relationship between logs and listeners, but by adding more kinds of data into the logs themselves. For outside critics of big data who had harped on the importance of context for making sense of aggregated interaction data, this operationalization of context—its addition to the logs rather then the recognition of it as something exterior to the logs—was perhaps an unanticipated response to critique (Seaver 2015).

Zero UI, in its attempt to “capture the attention of indifferent listeners,” represented a new kind of trap. This trap extended its threads out into the various data sources that users often do not even realize they are filling up with information about themselves, interacting unwittingly with a system designed precisely to capture those unwilling to interact. The design of the trap follows from the understanding of listeners primarily in terms of avidity and the consequent ideas about how less avid listeners behave. Recommendation in this mode can be understood as a system for increasing the avidity of listeners, for encouraging them to listen more than they might have and for facilitating interactions (even in the absence of intent to interact).[31] Where the musical identity advocated by Tom was an aggregation of interactions, the listener imagined by Zero UI was something more: a member of various statistical distributions, waiting for interactions to actualize their identity out of the potentials predicted by certain known features—among them, demographics.

This chapter began by describing the ascent of behavioral understandings of listeners and the users of computing products more generally. The behavioral data Peter suggested collecting fit well with the ethos of *Hooked* and Mike's desire to capture new users however possible, privileging "implicit" signals gathered from interactions. Avidity, whether understood as a kind of musical identity or a situational preference for certain kinds of interaction, echoed the concerns of behaviorism, in that it allowed engineers to interpret a wide range of activity and significance through the narrow measure of interaction frequency. However, unlike orthodox behaviorists, my interlocutors were more than willing to speculate on the lives of users only partially collected as logs of interaction events. Peter's desire to add to the user profile features like demographic information spoke to the expansiveness of big data logics of correlation beyond the narrowly behaviorist.

[31] This explanation for the use of implicit behavioral signals can sit alongside other explanations that focus on a broader discourse in which implicit behavioral signals are "truer," communicating more and more accurately than explicit statements. Mark Andrejevic describes this ideology, following Slavoj Žižek, as the "demise of symbolic efficiency," and the resulting attempt to glean advantage from patterns in data without special regard for their referents as "post-referentiality" (Andrejevic 2013).

The use of demographic data was a departure from the sensibilities that had shaped research and development in recommender systems since the 1990s. In their book *Word of Mouse*, John Riedl and Joe Konstan, two of the early figures in academic research on recommenders, argued that algorithmic personalization technologies were set to displace demographic marketing. "The urge to poll and classify is intoxicating" (Riedl and Konstan 2002, 109), they wrote. "The problem is, simple demographics don't begin to tell the story of individuals" (Riedl and Konstan 2002, 112). Where market demographics were coarse and biased, recommenders could target narrowly and respond to actual patterns in usage data, finding unexpected or newly emergent groups and allowing for users to change over time. This responsiveness was a moral imperative: they wrote, "Racial profiling and profiling your customers both spring from the same lazy, prejudiced philosophy" (Riedl and Konstan 2002, 113). "Think," they wrote, "about how much more people would step outside their demographic groups if they were not only permitted to, but *encouraged* to" (Riedl and Konstan 2002, 112). Among recommender systems researchers, it was often noted that adding demographic data did not improve the performance of systems, provided they had enough usage data.

However, while in pursuit of solutions to the cold start problem—in the construction of traps designed to catch listeners who had not yet interacted or would not interact enough—demographic information reappeared as a viable technique for drawing indifferent users into correlations with other users, capturing their attention and time, and locating them within a web of partial statistical connections. Peter and his colleagues produced a series of maps and data visualizations to display the most "distinctive" music for various demographic groups—states, genders, ages—calculating what was uniquely popular among women in their 30s or in Vermont. The implication was that these tendencies could provide new defaults for recommendation: without knowing anything else about a listener, these demographic tweaks might entice them to stay with a slightly higher probability than another default. That it seemed to go against some of the founding principles of algorithmic recommendation was of little concern to my corporate engineer interlocutors, who were willing to take whatever measures necessary to captivate users. When it comes to the large market of indifferent listeners, it is easier to catch them with a coarse probabilistic net than with precisely, individually targeted arrows.

Counter to the claims of some supporters and critics, these systems are not inherently atomizing or methodologically individualist (see, e.g., Galloway 2004, 111–113). Rather, they produce new collectivities, shifting among a variety of correlations that locate users as members of newly constructed and dynamic groups. These correlations take as their input a variety of signals which previously reflected distinct social epistemologies—behavioral data, demographic information, and so on—and they blend them together, without regard for the idea that these approaches reflect incompatible founding assumptions. At times, they evince the flaws of behaviorism—mistaking measurements for the thing being measured, neglecting the contextual specificities that make one click different from another—at other times, the flaws of demographic market segmentation—overestimating the coherence of a pre-defined group of people. But it should be noted that personal refutations of these approaches—"I do not listen like other men in their 20s, so it won't work on me"—misunderstand how these systems come to work: in the aggregate, eking out improvements that only manifest at large scales, and in conjunction with numerous other signals, which will be rebalanced at the first indication that a given signal is not working out.

These understandings of listeners grow alongside the development of algorithmic systems, mutually informing each other: the availability of particular kinds of data support particular ways of understanding users, while those understandings of users justified the design of systems influenced by them. The development of preferential technics is marked by these mutualities and feedback loops, through which some understandings are reinforced at the expense of others and integrated into circulatory infrastructures. Traps and the prey they are designed to catch are caught up in each other, and they proceed from partial understandings of what prey are like rather than from an objective account of their traits (however existent or knowable those traits may be). Logs of behavioral data and, increasingly, statistical tendencies are key elements in these captivation techniques which aim, as Franck Cochoy argued, to encircle and guess the trajectories of their targets (Cochoy 2007, 212).

The Ethics of Captivation

When I presented this argument—that we can profitably think of recommender systems as analogous to animal traps—to a room full of research scientists at an industry lab, one came up to me with a complaint: Certainly physical trapping was not the most adequate analogy to algorithmic recommendation, which, after all, does not literally trap the user in a box or end their life. A better model, he suggested, might be a kind of behavioral trap popular in folk tales and on television: the jar full of food that catches an animal's hand, which can't be removed until it lets go of the pickle or berry it's holding (TV

Tropes, n.d.; Kimmel and Hyman 1994). In this case, the scientist told me, the animal is trapped because it can't—he corrected himself—because it *won't* let go. He had slipped in the blurry middle in the understanding of captivation advanced in this chapter, between mental and physical captivation. "Won't" implies that the animal is responsible for its own continued trapping, while "can't" reaffirms the animal's helplessness in the grips of a trap that uses mental captivation for the purposes of physical captivation. The folk tales about this trap make the moral difference clear: greedy creatures pay the price for their pursuit of excess. In behavioral traps the trapper is, as it were, off the hook.

This space has been well-charted in Natasha Dow Schüll's expansive account of another psychologically tuned trap: the slot machine. She describes the various practices taken by slot machine makers to draw gamblers into the "zone," where they sit at machines and spend money for hours, stuck again in the blurry middle of mental and physical captivation. She quotes from a forum for gambling addicts, where people shared their experiences of captivation in front of the machine: "Something sinister was at work here, enticing 'normal' people into a snare"; "when I gamble, I feel like a rat in a trap" (Schüll 2012, 105). Although it seems unlikely that music recommenders would spur addiction—in spite of Mike's rhetorical flourishes from the beginning of this chapter and the claims of *Hooked*—the trap frame reminds us to pay attention to questions of consent and power in emergent technologies of captivation. It reminds us that technical systems exert a delegated agency and have the capacity to captivate people against their will.

These questions have been raised most loudly in regard to Zero UI-like systems, which glean scanty advantages from extensive data collection. Critics of big data question the voracious aggregation of any possible data, noting that for many, such aggregation poses risks, potentially violates privacy, and does not provide substantial benefit to outweigh these harms (Crawford and Schultz 2014; Tufekci 2014; Pasquale 2015). Peter and others working in this space worry about this perception—that data collection might come to be seen as “creepy” by users. But, the general consensus among my engineer interlocutors was that, over time, as data collection becomes normalized, this creepiness threshold will retreat, and users will consent to more and more. Critics would do well to attend to how the design of widespread data collection schemes is premised on particular understandings of users to be trapped: these understandings are what make data collection make sense to the people who construct it. If we want to push back against widespread data collection, it may be useful to push against this particular imagining of users first.

But there is more to the ethics of captivation than the general argument that captivation is bad—a denial of agency or the triumph of one actor’s agency over another. My questioner’s understanding of agency—a quality of individuals, which might be pitted against each other in struggles—was at odds with the understanding behind these theories of traps, which hold agency to be an effect of certain arrangements of persons and things (see, e.g. Gell 1998, *Art and Agency*; Barad 2007). The act of trapping is not just the application of a trapper’s agency against a prey animal. Rather the trap, as a dynamic sociotechnical scenario, is a tangle of agencies, aligning the tendencies of trapper, prey, and trap in time, as Gell writes, and requiring them all to function. The theory of captivation advanced by Gell and elaborated in this chapter takes it not principally as an object of critique, but rather as a generic feature of social life. People regularly try to captivate each other and seek out opportunities to be captivated, through art, technology or social interaction. Society and culture themselves can be considered as forms of mutual and superorganic capture.

Rey Chow notes that these theorizations of entrapment do not pay sufficient attention to the experience of being trapped—"of captivation as an experience that exceeds an ex post facto analysis of power relations" (Chow 2012, 48)—thus inviting the critique I received at my talk: what room does speaking of users as trapped animals leave for them to exert agency, to resist traps set for them, and doesn't this grant too much power and consequence to the decisions of developers qua trappers? Chow elaborates: "once caught, the prey's existence renders the trap more than just the elegant design understood from the sovereign command perspective of the hunter, who can henceforth no longer monopolize the terms of the interaction" (Chow 2012, 46). The objects of trappers' attention can and do resist entrapment, which after all depends on their participation — recall Otis Mason's account of animals becoming "more intellectual and wary" (Mason 1900, 660) in response to trappers' efforts, requiring ever more meticulous arrangements.

Yet for many, the language of trapping brings to mind the violence and imbalance of the relationship between predator and prey. Engineers resist the analogy to hunters—recommender systems provide much more benefit to their users than a dead-fall trap does to a Siberian sable. This suggests that we might look to another mode of captivation, which makes both the mutual benefit and sociotechnical nature of capture more evident. We can find such an example in the practices of pastoralism. As Tim Ingold observed in his study of Sámi reindeer pastoralists in northern Finland:

> Much of the knowledge concerning the behaviour of wild deer, used by hunters to trap the deer by deceit, is utilised by pastoralists to different ends: to achieve, in effect, a relationship of control through symbiosis. (Ingold 1974, 525)

Sámi pastoralists track the reindeer through winter and summer migrations, modulating their environment enough to keep them around: they provide emergency food, they light fires that keep mosquitoes away, they entice them into temporary, ever-changing enclosures. Pastoralism provides a way to understand the captivating practices of recommender system developers, who after all do not want to annihilate their prey, but rather want to keep them around and grow their numbers. It also provides an example of captivation in which the capturers play an involved, ongoing role, as the developers of recommender systems do, tending to their traps over time rather than setting them and lying in wait. Though Ingold emphasizes the pastoralists' control, later scholars of pastoralism note that this "control is rarely complete," and is rather "an ongoing exchange with the self-determined behaviour and preferences" of the pastoralist's herd (Reinert 2015; see also Reinert 2008). Pastoral enclosure is a kind of non-lethal, ongoing relationship aimed at growing the number of creatures enclosed through the careful social organization of animal and environment. "For the pastoralist, capital lies in the herd" (Ingold 1974, 526).

Pastoralism thus points to the importance of growing the numbers of captivated entities, a sense that already captured entities will cause their own numbers to grow, and it also points to an understanding of the trap as a kind of environment for the entities trapped. These concerns resonate with the concerns of music services: as mentioned earlier, their primary goal is to grow their userbase, and it is common Silicon Valley sense that users beget more users, making exponential growth a key signal of a startup's health to investors (Graham 2012).[32] In his writing on traps, Gell had suggest that traps "are lethal parodies of the animal's *umwelt*," or sensory environment; "Thus the rat that likes to poke around in narrow spaces has just such an attractive cavity prepared for its last, fateful foray into the dark" (Gell 1996, 27). In pastoralism, this shaping of environment happens at a wider scale and over longer durations, shaping not only the ultimate space in which prey will find itself but an environment in which prey can live and even flourish, albeit under terms broadly set by the pastoralist. Recommender systems are traps, but they are also *spaces*. As it turns out, this understanding of recommendation as the production of a space is popular among the developers of these systems as well. In the next chapter, I turn to how my interlocutors understood the space of recommendation and how these understandings work to mediate the questions of agency and control raised here.

[32] Like any analogy, this one is not absolute: where startups aim for exponential growth, with the end goal of a successful "exit" or acquisition by a larger company, ecological understandings of pastoralism note that herds cannot grow out of control without threatening balance in the system (see also Rappaport 1968). Where pastoralism has appeared to be a form of capitalism to some critics, this crucial distinction sets them apart (Ingold 1988).

CHAPTER 3:
PARKS AND RECOMMENDATION

"We say the map is different from the territory. But what is the territory?" – Gregory Bateson, *Steps to an Ecology of Mind* (1972: 460)

Music is a World

In 2014, Edward, an engineer working for a large music streaming service, presented a musical map to an audience at an American pop music conference. He began by narrating a journey, illustrated with musical examples: the formation of his own taste, from the soft folk of his childhood, through the transgressions of heavy metal, to the synthesizers of the 1980s, which broke open his musical horizons with the "transcendent crushing power and bewitching sparkle of robots." From there, he started to aggressively pursue new music and his taste broadened (he started to introduce his examples with their countries of origin): music from Scotland, Malawi, Australia. "We are now at the dawn of the age of infinitely connected music," he told his audience, referring to the growth of on-demand streaming services. Since his own musical horizons broke open, it has only become easier to hear music from around the world, and he has committed himself to listening to as much of it as possible.

This journey brought him to the present, where, among other things, he analyzed genres for an on-demand streaming service. Thanks to the emergence of large-catalog, on-demand streaming services, ordinary music listeners had 30 million songs available to them, from all over the world. This "isn't all the music in the world, but its way more than you'll ever have time to hear, so don't be an ass about it,"he added. This invocation of scale resonates with the accounts of information overload described in Chapter 1, but for Edward, the problem was not that too much music reached him, but that he could not reach enough music. Now, with inside access to large amounts of listening data and other information about music scraped from the web, this was the problem he had set to working on: facilitating music discovery. He remarked: "I sit at the center—as literally as there can be a literalness to this idea—of this universe of music as information."

Thus, he began another journey, taking his audience on a "tour of a map of the music-genre space," a wide-ranging, para-ethnomusicological ramble through musical styles from around the world, as seen from his seat at the "center." This music not only came from across the world — he would play his audience snippets of music from Japan, Croatia, and Zimbabwe—but also constituted a "world within itself" (Wonder 1976), with musical patterns cross-cutting geographical difference—from piano music whose "simplicity transcends cultures" to metal and techno and rap from unexpected places.

This trope, that music occupies or constitutes a space and recommendation enables journeys through it, is the object of this chapter. Edward was not alone in his characterization of music as a space, in which similar things were near each other and listening to new music was analogous to traveling. Throughout my fieldwork, I talked with graduate students working on the mathematics of vector spaces, in which music and listeners could be placed, I talked with hackers who envisioned their software projects as tour guides for taste, which was a kind of trajectory, and I talked with businesspeople who tried to understand and explain how the "music space" was like or unlike other kinds of cultural market spaces. This idea is much more general in the world of machine learning and beyond: “similarity spaces,” “vector spaces,” and the like are often the objects of machine learning algorithms, which try to correctly discriminate between different kinds of points in a space, or to map items into a new space such that their arrangement reflects their similarity. These understandings of space organize the work of algorithmic systems in their image, providing a frame in which to interpret algorithmic outputs, to make design choices, and to imagine what computational processes are doing.

Here, rather than taking on the question of spatialization in statistics and computing more generally, I focus on a narrower question that comes to bear on how music recommendation developers understand their work: What is the nature of the music space and the connections between items within it? My contention here is that music-as-space is central to the ways that the builders of recommender systems think about their work, and that in the diversity of ways these spaces are imagined and constructed, it is possible to see the connections between theories about culture and the technical systems they inform and are informed by. Similar efforts to discern and spatialize cultural patterns can be found in the history of anthropology, as can critiques about the correspondence between these patterns and other modes of spatialization, such as national or geographic boundaries. In this chapter, I bring these two traditions—anthropology and algorithmic recommendation—to bear on each other, demonstrating a shared set of concerns regarding the relationship between space, nature, invention, and control.

The map across which Edward traveled with his audience was a product of his job managing a genre ontology: the growing set of about 1,130 genres into which his company classified music. Where other music infomediaries used pre-defined genre hierarchies—all the jazz subgenres as branches of Jazz, every artist or song in its proper place on the tree—his genres were classified bottom-up, induced from patterns in online chatter about music and other signals. Drawing on these signals, the company produced measures of artist similarity—a number that, for every pair of artists, described how alike they were. From these similarity numbers, an algorithm could locate groups of artists that went together, according to common terms used to describe them or listening activity. Thus, the Norwegian symphonic black metal group Dimmu Borgir ended up clustered with Norwegian metal compatriots Gorgoroth and the American symphonic black metal band Wykked Wytch, while American pop artist Britney Spears ended up with her ex-boyfriend Justin Timberlake in "dance pop" and the cast of the movie *High School Musical* in "teen pop." These genres could shift over time, reflecting the changing dynamics of their underlying signals, and artists could be part of multiple genres at once.

The arrangement Edward decided on for his map arranged the genre names by their typical acoustic features: "The calibration is fuzzy," he wrote, "but in general down is more organic, up is more mechanical and electric; left is denser and more atmospheric, right is spikier and bouncier." Upon clicking a genre name, a user would be treated to a short, representative sample; another click would bring up a map of artists within that genre, again organized by their typical sound.

The result was a blob of names shaped roughly like Madagascar. In the bouncy and electric northeast were various forms of the dance genre House—Deep Tech House, Minimal Tech House, Acid House, Chicago House. If you traveled south along the east coast, you would pass through a few global dance musics — Baile Funk, Kuduro, Reggaeton—before coming upon a Hip Hop bay (the analogous location in Madagascar is called Antongil Bay), full of national variants — Canadian Hip Hop, Russian Hip Hop, Hip Hop Quebecois. Continuing south along Madagascar's highway 5, some diverse genres often collapsed into "world music" appeared—Highlife, Norteño, Mbalax, Malagasy Folk (the latter featuring Madagascar's own Rakotozafy, a renowned player of Malagasy zithers). In the spiky, organic southeast, the map broke apart, with a coastline of Jazz and Blues looking out on to an archipelago of spoken-word recordings—Poetry, Guidance, Oratory, Drama. The dense, atmospheric west coast was dominated by Rock and Metal—from Dark Hardcore in the north to Black Sludge in the south. The two coasts were separated by a ridge of pop genres—from Japanese Shibuya-Kei to French Yé-Yé. At the bottom of the island, a Classical peninsula jutted out—Opera, Carnatic, Polyphony, Concert Piano.

These unusual genre designations were not tied to musicological (Dahlhaus 1983) or sociological (Lena 2012) definitions of genre, but were intentionally diverse, referring to types of instrument ("Cello"), high-level industry categories ("Pop"), specific local traditions ("Forró"), or clusters that almost, but not quite, resembled existing genres ("Deep Cello"). They drew on the ordinary practices of listeners instead of the classifying schemes of experts. The bottom-up ethos of Edward's clustering coincided with his ideas about genre. He would not insist that his groupings were correct or objective, but rather that they were useful, and no worse than many other potential starting points for defining musical genre. “This process isn’t entirely accurate or precise,” he wrote, “but music isn’t either, and they both often seem to work.” Rather than claiming objectivity for his map, he offered it as a tool for exploration: “The purpose of the map, as it is for the genres, isn’t to end arguments but to invite exploration of music.” For Edward, this tentativeness and pragmatism—where genre classification was not a top-down imposition of order but an inherently malleable and emergent thing, constantly changing with the acquisition of new data and at the self-aware caprice of its coders—was a virtue.

Edward would occasionally gloat on his Twitter account that the audio classifier had so successfully grouped his genres without being told about their relationships to each other: Hip Hop cove was full of hip hop genres not because they were called “hip hop,” but because the computer thought they sounded alike. “If someone tells you ‘algorithms’ don’t understand music, show them this and take their headphones away,” he tweeted, posting an image of regional hip hop genres stacked neatly on top of each other. The neat stack indicated some agreement between the signals that fed into algorithms that generated the genre clusters and the audio analysis that had arranged them in space — two distinct techniques that seemed to verify each other’s outputs. In publicity materials, these two sides of his company—the one that analyzed online chatter using natural language processing and the one that analyzed musical sound using machine listening — were called the “cultural” and “objective” sides, respectively.

If music is a world, then music recommendation is a kind of guided navigation through it. “For this new world to be appreciable,” Edward told his audience at the conference, “we have to find ways to map this space and then build machines to take you through it along interesting paths. Right now the paths are rough and the tour buses are rickety, but we’re starting to find our way.” The two journeys he narrated in his talk reflected a dominant way of being in this musical space: in transit, encountering novelty and difference. The resonance between these concerns and those of anthropology is plain, rooted in an interest in the new, the exotic (or the “new-to-me”), and in traveling. “Follow any path, no matter how strange and barren it seems, and you’ll end up in secluded spaces with a hundred bands who’ve lived there for years, reconstructing the music world in methodically- and idiosyncratically-altered miniature, as in Austrian hip-hop, Slovak pop, microhouse or Latin metal.”

Edward’s tour of a map of a space that is both a world and a universe demonstrated a productive confusion between maps and territories: is the world of music a space to be mapped or is the map of music itself a space? Does music constitute the space or occupy it? The often vivid metaphors through which developers explain and apprehend the spaces they work in and on are revealing and influential regarding their answers to these questions and the decisions they make in the course of their work. When I showed his map to an executive at a competing infomediary company, known for its hierarchical genre model, the executive mocked the idea: “Those hipsters,” he said, “they don’t want to admit that genres *are* hierarchical, and they want to pretend that anything can be related to anything else.” These understandings of space tend to rely on a set of metaphors that mediate between two extreme positions regarding the nature of the music space: on the one side, that it is something *discovered* by engineers through the analysis of their data (making it something like a territory), and on the other, that it is something *constructed* by them (making it something like a map). This tension is familiar from the philosophy of science and science and technology studies: where discovery implies the uncovering of a natural order and the work of objectivity, construction implies the operation of cultural bias and the exercise of control. It can already be seen in Edward's account of his own work, which oscillates between confidence in the performance of algorithms and emphasis of the arbitrary terms of its presentation; it also hums in the peculiar distinction his company draws between its "cultural" data and its "objective" data.

The question of cultural space is not new to the developers of recommender systems. For the market researchers these systems elaborate on, the language of "space" has been common for some time. And although contemporary anthropologists might not recognize these formalized spatial analyses of culture, they also have precedents in the history of anthropology. In the rest of this chapter, I trace out the understandings of cultural space that preceded my interlocutors' ideas about space and my own, noting resonances between the history of anthropology and algorithmic recommendation. These ideas about space are centrally concerned with connections among persons and things, and they provide a variety of ways to think about the relationship between space as a construction an as an objective fact.

The Synaptic Function

In his 1976 book *Culture and Practical Reason*, Marshall Sahlins drew a provocative comparison between anthropologists and marketers.[33] An anthropologist, he wrote,

> acts in something of the same way as a market researcher, an advertising agent, or a fashion designer, unflattering as the comparison may be. For these hucksters of the symbol do not create de novo. In the nervous system of the American economy, theirs is the synaptic function. It is their role to be sensitive to the latent correspondences in the cultural order whose conjunction in a product-symbol may spell mercantile success. (Sahlins 1976, 217)

While anthropologists like Victor Turner limned the significance of colors in Ndembu ritual, Madison Avenue creatives drew together colors and images to advertise detergent or fast-food hamburgers to American families. In both cases, according to Sahlins, the analyst was concerned with finding "latent correspondences in the cultural order"—between redness and a Central African river, or between the image of a forest and the notion of "cleanliness." This sensitivity to pattern and style—the way certain symbols, practices, and techniques "go together" for certain people—has characterized anthropological thought throughout its history (e.g. Graebner1911; Benedict 1934; see Wilf 2013).

Marketing, Sahlins claimed, was a form of "bourgeois totemism" (Sahlins 1976, 178), classifying groups of people through their correspondences with groups of things. Rather than sorting into "raven" and "wolf" moieties, members of bourgeois society might sort into "Pepsi" and "Coke," replacing the natural schemas of conventional totemic organization with man-made ones. Sahlins is using "totemism" here in a

---

33 Parts of this section and the next two have been adapted from (Seaver 2012).

restricted technical sense: in Lévi-Strauss's terms: "a classificatory device whereby discrete elements of the external world are associated with discrete elements of the social world" (Lévi-Strauss 1963, 7; cited in Lien 1997, 239). Following Lévi-Strauss in taking totemism to be a mode of anthropological analysis rather than a coherent practice encountered during fieldwork (Lévi-Strauss 1971), totemism provides an archetype for classifying *persons* through their correspondence with *things*. This turns out to be significant for understanding the operations of recommender systems, which, although they depart from previous forms of classification in many ways, maintain this basic totemic operation. Recommenders know what people are like (and *who* they are like) from knowing *what* they like. But if a basic understanding of totemism only allows for a few groups and a few objects to obtain a one-to-one correspondence, these new modes of the synaptic function extend the diamantine structure of Lévi-Strauss's totemic operator (Lévi-Strauss 1966, 152) into multiple dimensions.

In her ethnography of a Norwegian marketing firm, Marianne Lien noted the constructive aspect of marketing work—marketers did not simply identify correspondences between groups, but actively created them: she recounts the development of a line of frozen pizzas:

> In spring 1992, Viking Foods manufactured six pizza products on the Norwegian market. [...] The emergence of the present product range is a result of careful considerations of the characteristics of real and potential target groups. (Lien 1997, 171–2)

The resulting line of pizzas was designed to map to groups defined by their putative relationship to pizza: Pizza Superiora was "the people's pizza," for a general audience.

Pizza Romano was more expensive and had “a distinctive flavor and character,” targeted “a more adult and selective audience.” Pizza Preciosa had a wholemeal crust and vegetable topping, and it was aimed at “women aged 15-40 focusing on health, body and appearance. Vegetarians” (Lien 1997, 171–2).

Where Sahlins claimed that marketers and anthropologists did not create “de novo,” but were rather sensitive to actually existing correspondences in the cultural order, Lien demonstrated how marketers summoned market segments into existence as they described them, producing the conditions in which potential pizza fans might find themselves aligned to one kind of pizza and not another. This interactive dynamic, which might be called “performativity,” a “looping effect” (Hacking 1991), or simply “construction” is a common concern among social scientists. By organizing particular arrangements of persons and things, these hucksters of the symbol create cultural orders as much as they discover them.

This productive aspect of the synaptic function means that the terms through which the cultural order is imagined—as a nervous system with synapses or a diamond logical structure, for example—become quite important, as they guide the work of construction. And although social scientists may be the most explicit in defining their cultural imaginaries *as* imaginaries (think of Geertz’s “webs of significance” or Benedict’s “patterns of culture”), it is the imaginaries of other “hucksters of the symbol” that have the most potential for broad influence, dominating popular discourse or being built into influential infrastructures. Yet Sahlins was right to note a latent correspondence between the work of people who look for latent correspondences. We can trace connections between various modes of the synaptic function: between marketing and the recommender systems that claim to be its successor, and between these commercial efforts and those taking place in the academy.

His Bag is People

Although Sahlins was being cheeky, his contention that market research is effectively anthropological has been borne out by the growing ranks of anthropologists and ethnographers applying their work to commercial purposes (Cefkin 2009; Nilsson 2013; see also the Ethnography Praxis in Industry Conference). Indeed, by the time *Culture and Practical Reason* was published in 1976, anthropologists were already being cited in the pages of the *Journal of Marketing Research*.

One of those anthropologists was Volney Stefflre, a professor at the recently founded UC Irvine School of Social Sciences. A 1969 profile in *Orange County Illustrated* with the groovy title "Volney Stefflre: His Bag is People" described Stefflre as "half anthropologist, half college professor, and half business executive. If this adds up to one-and-a-half, it is because Volney Stefflre is an oversized man, physically, intellectually, and enthusiastically" (Van Deusen 1969, 31). While Stefflre's colleagues modeled kinship terms, occupational prestige, and ethnobotanical classification in societies around the world, Stefflre worked closer to home—studying the responses of suburban residents of Orange County to consumer goods and advertisements.

The School of Social Sciences had no departments: under the leadership of organizational scientist Jim March, it was expected that anthropologists, psychologists, economists, and sociologists would work together, with "substantial disrespect for traditional disciplinary identifications" (Kavanagh 2010, 8). The various scholars were loosely united by a commitment to formal and quantitative methods. Taking part in the post-war "cognitive revolution," UCI became a leading site for the development of mathematical and cognitive anthropology, and members of the school of social sciences pioneered new techniques for multidimensional scaling and the use of computers in formal analysis of "culture" as a shared patterning of mental models (e.g. Romney, Shepard, and Nerlove 1972; Burton 1973).

Using data collection and analysis techniques from linguistic and cognitive anthropology, Stefflre presented his subjects with fabricated advertisements, lists of snack foods, samples of toilet paper, and bottled drinks of different colors. In a series of experiments, he determined which colors, shapes, and images people associated with abstract concepts like "cleanliness" or "health"; he analyzed the correlation between snack foods like pretzels or ham sandwiches and use cases like "after a party" or "for breakfast." From this data, Stefflre produced what he called "market structure analyses": 2- or 3-dimensional plots that arrayed existing products in space according to their computed similarity. These plots represented the latent structure of the market, revealing submerged correspondences in the cultural order of consumer products.

This work had two audiences: in *American Behavioral Scientist*, he published on "people's behavior toward new objects and events" (Stefflre 1965), working on a more general theory of how different groups of people respond differently to new things. From his market research consultancy, he advised companies to create a new kind of division that he called a "New Products and New Enterprises Group." In a book he wrote advocating the idea, he described the mission of the group, which operated in the space defined by market structure analyses:

> A New Products and New Enterprises Group can be seen as a greased hole in the institutional and psychological wall that separates *what exists* from *what could, but does not yet, exist*. The wall—which is built of customs, institutions and people—prohibits the appearance of technologically and economically attractive new product alternatives that consumers desire and are willing to pay money for. (Stefflre 1971, 3-29; emphasis in original)

According to Stefflre, the latent structure of product-symbols (to use Sahlins' term) was a kind of mental-cultural architecture, which, once understood, might be reconfigured and monetized. Empty spaces in the structure reflected products that didn't exist yet and which companies, interpreting the space, should try to create. Stefflre, whom colleagues remember as a brusque and immoderate personality, vividly described this task as "a suicide mission—Kamikaze pilots—trying to bring into existence something that was not before—at a cost to themselves of years of their life and the tortures of the damned" (Stefflre 1971, 43-45). To this end, he started a series of consultancies and bought a local supermarket in which to conduct ethnographic experiments in vivo—what fellow formalist anthropologists would come to call "white room ethnography" (Metzger 1963; Black 1963) The goal of his cultural analysis was not only to passively perform the synaptic function, but to orient this knowledge toward the making of new products, which could then take their places in the spatialized organization of product-symbols.

Stefflre, a paradigmatic example of Sahlins' "huckster of the symbol," embodied the connection between marketing and anthropology that Sahlins had only suggested. Yet he did not think of his work as simply representing cultural space. His purpose was to intervene at the boundary between the actual and the potential (and, in his academic work, to understand how people behaved there), and his interventions were shaped by his architectural understanding of the product space. The order of product-symbols (with the corporate structures that maintained it) was a building to be remodeled, to be broken open and crashed into. His understanding of culture as space and his understanding of what could be done in that space went together.

On the Internet, there's no excuse for not personalizing

Over the next several decades, the computational techniques for multidimensional scaling developed by Stefflre and his colleagues would lose favor among anthropologists, as formalism was spun out of sociocultural anthropology by successive symbolic, interpretive, and reflexive turns. However, the broader movement in which they were participating—a turn to formal, computational methods inspired by cognitive understandings of human action—continued to grow in other fields.[34] These techniques for scaling persons and things into cultural spaces would evolve into contemporary data mining and recommendation practices (Desrosieres 2012). Though anthropologists who encounter these techniques may be taken aback by what seems to be a novel quantitative encroachment on their area of expertise, their lineages trace back to the same moment of post-war social scientific formalism and the "cognitive revolution" (Seaver 2015).

[34] At Stefflre's home institution, these endeavors persisted in a new academic program formed to house them: the Institute for Mathematical Behavioral Sciences.

Where Stefflre understood cultural space as a kind of architecture, which corporations might remodel, the cultural space imagined by developers of recommender systems was more dynamic and less stable. In their 2002 book *Word of Mouse: The Marketing Power of Collaborative Filtering*, two of the founders of academic recommendation research promised that recommender systems would upend the marketing world, understanding consumers not as members of demographic groups (those coarse totemic divisions Sahlins referred to), but as individuals. Though "the urge to poll and classify is intoxicating" (Riedl and Konstan 2002, 109), they wrote, "simple demographics don't begin to tell the story of individuals" (Riedl and Konstan 2002, 112). With the advent of online retail and new technologies for tracking the activity of customers, marketers could begin to follow these individual behavioral histories, targeting users not through generic demographic profiles, but with personally tailored recommendations: "On the Internet, there's no excuse for not personalizing" (Riedl and Konstan 2002, 112).

The book's cover promoted algorithmic recommendation as the equivalent of ESP for Sahlins's synaptic function: "Know what your customers want even before *they* do." Below that slogan, a cheery and diverse crowd of customers waved from inside a computer monitor, apparently pleased by this technological breakthrough in taste prediction. This group represented the "collaborators" of collaborative filtering—the users whose aggregated activity could be algorithmically mined to predict each other's preferences. Although these users do not know or communicate directly with one another, through the algorithm they are made collaborators—a computationally arranged aggregate of taste-bearing individuals.

If the market segment is the paradigmatic collective form of demographic marketing, this group of users inside the monitor might be the paradigmatic form of algorithmic recommendation. "Think about how much more people would step outside their demographic groups if they were not only permitted to, but *encouraged* to," wrote Riedl and Konstan (112). The friendly crowd on the cover appears to cut across traditional demographic categories of race, gender, and age, and Riedl and Konstan argued that tastes and preferences similarly cut across these conventional lines. Unhindered by externally imposed categories, these individuals are free to follow their own preferences, facilitated by the suggestions of the recommender, which could even encourage users to broaden their horizons by locating items that the broad brush of market segmentation would miss.

In this view, algorithmic recommendation represents a step in the ever-finer division of audiences and markets. If demographic marketing is "bourgeois totemism" as Sahlins suggests, then the progressive thin-slicing of market segments (e.g. the invention of "tweens") can be seen as an increasingly specific totemic operation, with individually targeted personalization as an endpoint. Arguments against algorithmic personalization often take this apparent methodological individualism as an object of critique (e.g. Turow 2011). However, algorithmic recommendation is not simply a higher-resolution representation of a market—a more precise picture of atomistic individuals that does away with the need for larger-scale approximations like market segments. Rather, it is another mode of the synaptic function—another technique for making and interpreting correspondences between persons and things, another way of organizing collective forms. Collaborative filters algorithmically rearticulate the relationship between individual and aggregate traits: the positions of individuals vis-à-vis groups can change continuously. John Cheney-Lippold has recently described this kind of algorithmic interpellation as "soft biopolitics" (Cheney-Lippold 2011), a shifting mode of categorization that necessitates a reappraisal of the models of power and taxonomy in Foucauldian biopolitics. Collaborative filtering does not merely privilege individuals over broader demographic categories; it reinstalls them into an algorithmically tuned collective. Collaborative filtering atomizes the totemic function, sweeping users up into temporary groups based on their partial connections and similarities to others.

Making Similarities in the Matrix

> "Have you ever wondered what you look like to Amazon? Here is the cold, hard truth: You are a very long row of numbers in a very, very large table." (Konstan and Riedl 2012)

Collaborative filtering is the archetypal form of algorithmic recommendation. As described earlier, the first developers of recommender systems envisioned these techniques as a kind of information filter that could be shared (Goldberg et al. 1992). Instead of creating one's own filters for email or newsgroups, one could share filters with others. Soon, this sharing and production of filters was automated, relying on profiles of users and items to "automate word of mouth" (Shardanad and Maes 1995). Although, as discussed throughout this dissertation, a number of other recommender techniques exist, the basic collaborative filter remains popular, both as a component of commercial algorithmic recommender systems, and as a toy project for computer science classes. In early 2012, I sat in on two such courses: an undergraduate course on building collaborative filters at UC Irvine and the popular online Machine Learning class on Coursera, taught by Stanford professor Andrew Ng. In these courses, collaborative filters proved useful for demonstrating the applications of algorithms and for providing intuitive outputs that were easy for students to parse.

A collaborative filter is essentially concerned with a table—a matrix with items along one side, users along the other, and ratings in the cells at their intersections. This table is mostly empty (or "sparse"), since most users will not have rated most items. The work of the collaborative filtering algorithm, as it is typically stated, is to predict what values will show up in the empty spaces of the table. As I learned in these courses, these predictions are typically made by locating users and items in a space defined by the numbers in the table (or in derivatives of it). These "similarity spaces" are constructed such that similar items and users end up near each other, making future calculations of similarity a matter of calculating distance—to recommend ten items similar to a given item, once simply needs to find out which ten items are closest in the space. The production of space is critical to the functioning of recommender systems more generally: while collaborative filters locate entities in space on the basis of ratings patterns, any other data might be used instead. Here, I outline in more detail how collaborative filter in particular works to locate users and items in space.

In the classroom at UCI, the instructor puts up a table with one of the rows marked "Alice," and the two cells in the row filled with "Dim1" and "Dim2"—her ratings for two items. "We've represented Alice with two numbers, so I can make a two-dimensional plot, and I can locate Alice," he says. Imagine a Cartesian coordinate system: if Alice rated two items 3 and 4 stars, she can be represented as a point at (3,4), with each axis representing one item. If another user, Bob, rated those items 1 and 5 stars, he would be a point at (1,5). Adding a third item could result in a third dimension, and although it is harder to visualize after three dimensions, an ever-growing list of items would result in an ever-growing number of dimensions in which users could be located as points. The resulting distribution of points would place users who had given similar ratings near each other.

It doesn't particularly matter whether the table is interpreted as a set of coordinates: to the computer, it is all the same. Distance calculations work on numerical inputs and do not mind the human-challenging jump from locating points in three dimensions to locating them in forty. (However, as I describe later in this chapter, such calculations soon veer away from human intuitions about space and distance.) To the humans building these systems, however, spatial explanations for what the computer is doing are central to understanding and explaining them. The "very long row of numbers" referred to above is commonly known as the "user vector," a set of numbers that locate a user in a space. Although there are a variety of mathematical kinds of space and ways of calculating distances (i.e. similarities), ordinary talk about recommender systems tends to blur and slide among them. This spatial common sense precedes any particular implementation, and my interlocutors regularly invoked spaces to frame their discussions of their work: "Everything lies in a space," as one grad student told me at a conference.

Now, imagine a user hasn't rated something — she can be located on two of the axes, but has no position on the third. Recommender systems use a variety of mathematical techniques to predict a likely location along that third axis, based on the distribution of existing points. If user ratings aren't random noise, but have some pattern, we would expect to see some clusters of points in our ratings space, and we would guess that our missing data point will fall somewhere in that cluster. Thus, at any given time, the matrix is in an anticipatory flux: new ratings from users arrive constantly, displacing their predicted values and shifting the others. This filling process is the signature action within the matrix—blank values are replaced by predictions, which are then replaced by actual ratings. Progress from emptiness, through prediction, to actualization makes the matrix a proleptic social representation, holding simultaneously a record of past correspondences between persons and things and the anticipation of future ones.

As I've described it, this is a "memory-based" collaborative filter: we use every data point in the table to try to locate users in a space, as we assume that missing values on one axis can be guessed by looking at a user's "nearest neighbors" on other axes. It is also a "user-user" system, calculating similarities among users to provide recommendations. It would also be possible to locate the items in a coordinate systems defined by users: If an item had a 4-star rating from Alice and a 2-star rating from Bob, then it would be located at (4,2) in the Alice-Bob coordinate system. This style of recommendation is called "item-item," because it calculates similarities among items rather than among users.

As Paul, a graduate student researching recommender systems, told me during a poster session at the RecSys conference in Dublin, these different techniques introduced different kinds of biases into the recommendations, and their performance often depended on the kind of data available to them. For very large catalogs, item-item recommenders tended to work best. Paul was trying to produce "engineering principles" for developing algorithmic systems—currently, he said, people would arbitrarily try many configurations, evaluate them against some accuracy metric, and then pick which one outperformed the others. His goal was to characterize the various approaches' strengths and weakness, so that developers might make informed choices among the options available to them.

The "ratings" at the intersections of users and items do not have to be explicit: music recommenders, for example, typically rely on "implicit ratings." Rather than asking users to rate a song on a 1-5 scale, the recommender will interpret certain actions as a kind of rating. Listening to a song all the way through, or repeatedly, may count as a positive rating, while skipping a song before the end may count as a negative rating. Implicit ratings are difficult to interpret (there are many reasons a user might make it to the end of a song they don't like—such as leaving the room—or skip a song they do like—because it was inappropriate for the context they were in at the moment), but because they require no active user effort to generate, they provide much more data than explicit ratings. Today, commercial systems that use explicit ratings will typically also incorporate implicit ratings, adding them to user vectors as more dimensions

Because these vectors can become extremely large, and a memory-based filter might have to deal with hundreds of thousands of entities, similarity calculations rapidly become computationally expensive. “Model-based” collaborative filters aim to represent items and users with a smaller vectors, referring not to specific rating interactions, but to statistical properties of those items and users in relation to each other. This is known as “dimensionality reduction,” and techniques for doing it are a key focus of research in collaborative filtering.

One of the more popular techniques for dimensionality reduction is called “singular value decomposition,” or SVD. This technique gained popularity with its introduction into the Netflix challenge, and it is a means for “decomposing” a matrix into “latent factors” along which users and items vary. Back in the undergraduate class at UCI, the instructor guides us through an implementation of SVD in the computer program Octave. We input a ratings matrix into the program and tell it to perform a “matrix factorization,” representing our table of ratings as two tables that capture the features of users and items in a new coordinate system. This new coordinate system is defined by the latent factors—the statistical regularities—of the ratings matrix. In it, users and items are not located by sets of ratings, but by sets of shared tendencies.

In a simplified plot on the classroom screen, the UCI instructor shows hypothetical movies, organized so that similar ones are near each other: “Magically, we’ve measured two secret things about these movies,” he says, referring to the two numbers that represent their new locations. In the example, they seem to have separated into action movies and romances, and the users have also spread across the space. A student is confused about how the algorithm extracts these genres from the ratings and the instructor says: “It seems like magic, I know. All the algorithm is doing is trying to reduce this error”—the error being the difference between the latent factors, which are an approximation, and the original data. That the movies appear to have sorted out by genre is validation that the latent factors represent something meaningful to humans. These factors don’t have to be interpretable to be useful in making recommendations, as I learned later (see the discussion of interpretation in chapter 4), but it helps, especially for explaining to students.

The production of this “feature space,” derived from the ratings that join users to items and within which both can be placed, is a critical step in algorithmic recommendation. Its didactic representation as a pair of easily interpretable axes is common: in an article written by Yehuda Koren, the winner of the Netflix Prize, he illustrates his method by distributing a set of hypothetical users and real movies across a coordinate system defined by “serious”/“escapist” and “geared toward males”/“geared toward females” (Koren 2008). In this coordinate system the corners are filled by *Braveheart* (serious, geared toward males), *Dumb and Dumber* (escapist, geared toward males), *The Princess Diaries* (escapist, geared toward females) and *The Color Purple* (serious, geared toward females).

People are distributed around this space according to their preferences, and Koren is careful to include a man on the “geared toward females” end of the spectrum, emphasizing the fluidity with which people can move through the feature space. Gendered axes are common in these didactic illustrations, reflecting a common sense about how taste and cultural objects are presumed to vary together. The recommender systems diagram also makes evident the role of the word "like" in "Users like you liked items like this": preference and similarity are collapsed in this coordinate system, where "being like" and "liking" have been equated. You may not like the same things as the rest of your demographic group, but you probably will share preferences with your "nearest neighbors" in the abstract cartography of collaborative filtering.

For readers accustomed to quantitative methods that remain more popular in sociology than in anthropology, the matrix factorization described above will be familiar. The method, and the resulting plots, bear a striking resemblance to Bourdieu's correspondence analysis plots from *Distinction*, which similarly locate persons and cultural objects in a shared feature space, derived from their affinities. Bourdieu's interpretation of his axes underwrote his theory of economic and cultural capital, across which people might move over their lives and in which certain groups of people (such as professions) might be located. But while the technical methods of algorithmic recommendation bore direct connections to earlier methods from the quantitative social sciences (Desrosières 2012), their spatial understandings of cultural variation also reflected a way of thinking about difference that was growing in popularity among their anthropological contemporaries.

Connection and Contiguity

Riedl, Konstan, and their colleagues in the early days of algorithmic recommendation identified the central flaw of demographic marketing as its reliance on rigid boundaries: in practice, individuals change, their preferences moving across supposed boundaries, and approaches to marketing that presume otherwise will fail and reinforce damaging stereotypes. This attitude toward cultural space and the motion of persons within it finds a surprising resonance in the 1990s writing of anthropologists on space and culture.

"Representations of space in the social sciences are remarkably dependent on images of break, rupture, and disjunction," Akhil Gupta and James Ferguson wrote in 1992 (Gupta and Ferguson 1992, 6). Anthropological space in particular, according to Gupta and Ferguson, had long been defined by discontinuities: the discipline was founded on the study of the remote and cultural difference was essentially premised on geographical distance. Thus "the distinctiveness of societies, nations, and cultures is based upon a seemingly unproblematic division of space, on the fact that they occupy 'naturally' discontinuous spaces" and "space itself," they wrote, "becomes a kind of neutral grid on which cultural difference, historical memory, and societal organization are inscribed" (Gupta and Ferguson 1992, 6, 7). To find another culture, you simply had to go to another place.

But, Gupta and Ferguson argued, the stacked differences of space and place, society and culture, seemed to be slipping out of alignment. The illusion that identities coincided with borders became untenable. Cultural difference manifested within specific localities, drawing into question ideas about cultures as unified wholes; people occupied the borderlands, crossing between supposedly discrete spaces and societies, drawing their definitions into question (Anzaldúa 1987); hybrid postcolonial cultures similarly problematized simple geographic maps of cultural difference; and the growth of global cultural and economic interconnectedness that Fredric Jameson called "postmodern hyperspace" (Gupta and Ferguson 1992, 8; Jameson 1991) seemed to short-circuit any attempt to isolate and describe a "culture." The convenient fiction of the isolated field site, contiguous with its geographical borders, could hardly be maintained when its residents watched American television and made products for European markets. While historical developments in communication and commerce ate away at the discreteness of anthropological field sites, they revealed that such discreteness was never as deep as it had seemed. It was a spatial trick, discontinuities borrowed from maps and fixed onto people whose samizdat continuities continued to flow across borders.[35]

For Gupta and Ferguson, the question was how to make sense of cultural difference when it could no longer be imagined as an outcome of spatial arrangement. If everything is potentially connected, then the issue becomes how cultural difference is *produced* from promiscuous flows and how space and place, those apparent causes, are actually effects.[36]

> Physical location and physical territory, for so long the only grid on which cultural difference could be mapped, need to be replaced by multiple grids that enable us to see that connection and contiguity—more generally the representation of territory—vary considerably by factors such as class, gender,

[35] This situation precipitated the very influential model of multi-sited fieldwork advanced by George Marcus in "Ethnography in/of the World System" (Marcus 1995) and conducted for this project.
[36] See also Boellstorff's "archipelagic" analysis of identity and geography in *The Gay Archipelago*, which goes beyond the vague declaration of "fluidity" to describe the constellated structure of apparently border-crossing identities in practice: "like constellations, archipelagos are networks, constituted through lines of connection" (Boellstorff 2005, 16).

> race, and sexuality, and are differentially available to those in different locations in the field of power. (Gupta and Ferguson 1992, 20)

For would-be analysts of culture, the displacement of geographical space by other modes of connection and contiguity required new attention to the politics of difference as they manifested along other axes. But these other spatializations did not do away with the organizing power of spatial metaphors. Rather, they dislocated the spatial from the geographic. Fields of power are still spaces, and it is telling that Gupta and Ferguson made the jump from geographical space to the field of power via the mathematical figure of the "grid." Even (and perhaps especially) in "postmodern hyperspace," math provides a common currency for trading among various sorts of difference. Though the mathematization of the social is something anthropologists tend to associate with outmoded post-war formalisms like ethnoscience or the neo-positivism of "data science," mathematical imaginaries persist, as Kath Weston has argued, even in "avant-garde social science metaphors: borders, lines, intersections, levels, scales, points, grids, and of course the 'trans' that introduces transverse and transept as well as transnational" (Weston 2008, 133).

Although Riedl and Konstan described their work on recommender systems as a turn from groups to individuals, the actual practices of recommendation bore more resemblance to the novel grids of difference described by Gupta and Ferguson. In their reference to "postmodern hyperspace," Gupta and Ferguson shared with Riedl and Konstan a common sense that the rise of large-scale network communication technologies would facilitate the production (and recognition) of new spaces of difference: for the former, these spaces were fields of power, for the latter, they were "taste spaces." Where the former emphasized the persistence of constraint, the latter suggested a novel freedom. While the growth of the internet inspired many inside and outside of the academy to suggest that cyberspace represented a break from existing structures of difference, geographical and otherwise (e.g. Barlow 1996), it became the work of critical scholars in the humanities and social sciences (e.g. Markham 1998; Chan 2014) to document the persistence of connections to previously existing systems of power and geography.

On Edward's genre map, the persistence of geography was most evident in the names of genres. Although music's contemporary availability made "teleportation" (as he glossed it) to different places possible, those places' difference was nonetheless often predicated on their geographical distinctness. For Edward, travel in the music space was closely analogous to travel in geographical space: some genre names, like "Slovak Hip Hop" or "Vietnamese Pop," identified musical styles with their nations of origin (reflecting the global circulation of dominant styles of music), while others referred to regional styles, like South Asian "Qawwali" or South African "Kwaito." The new spaces of music online were still shaped by the geography that had long shaped structures of musical difference.

Space and Choice, or the Curse of Dimensionality

The culturally laden nature of classification schemes has been well established by anthropological and STS studies of classifying practices (Durkheim and Mauss 1963[1903]; Bloor 1982; Douglas 1966; Bowker and Star 2000), and Edward's genre typology provides another example of how this cultural ladenness plays out. In clustering systems which work “automatically,” the resulting clusters are still identified and named, and this process provided a moment of choice and interpretation. Sometimes, a cluster would be filled dominated by an existing genre name, regularly appearing in the metadata—like “Country” —suggesting an obvious name for the cluster. Other times, a name would be more elusive, requiring Edward to pick which among a variety of names to use, emphasizing one aspect of a genre over another: such was the case with “Indie R&B,” the name Edward eventually assigned to an emerging style of R&B notable for the whiteness of its audience relative to conventional R&B. Critics had made several attempts to name the genre, including “PBR&B,” a portmanteau of R&B and Pabst Blue Ribbon, a beer brand associated with the same audience (Harvey 2013). For a while, Edward had chosen to go with “R-neg-B,” a name that emphasize the style’s typical negativity (Wilson 2011), but eventually he changed the name to “Indie R&B,” a tag with less implicit editorializing about the genre’s audience and affect, though it maintained the racialized understanding of genre which in which “indie” signified whiteness and “R&B” blackness.

Edward’s genre names emerged from an interpretive process, but so did the operations on either side of them: the algorithms that generated clusters would have been tuned according to the interpretability of their output, and the terms by which they were arranged in space were the result of another interpretive process. Interpretability is not a feature of outputs, but of the relationship between outputs and interpreters (see the section on interpretability in the next chapter). Edward had chosen to organize the names along organic/mechanical and atmospheric/bouncy axes, but these were only two options among many. “In fact,” he wrote, “we can build an almost infinite set of views of the world.” Eventually he would vary the color of the genre titles to indicate a third dimension of difference (which one, he wouldn’t say, preferring color to be interpreted as a vague feeling, rather than as an objective metric). If music was now “infinitely connected,” it followed that there were as many possible maps as modes of connection. Choices among maps were a matter of pragmatics rather than a scientistic pursuit for the ideal representation—they were not final statements on the order of things, but rather tools for exploring that could readily be altered, replaced, or exchanged in pursuit of other goals. For those like Edward, with access to the tools, training, and data, such maps were essentially malleable.

For Gupta and Ferguson, the recognition of new spatial perspectives did not displace old modes of difference, nor did they displace each other. Upon recognizing other forms of difference as kinds of “spaces,” they multiply, leading to intersectional identities in the midst of many-dimensional spaces (Strathern 1991). For music and other classificatory problems, this meant that objects could be located within a nearly infinite set of shifting axes of difference. This apparent freedom posed a challenge, however: as coders located music in more and more dimensions, computers had a harder time identifying salient distinctions and connections—a problem researchers called “the curse of dimensionality.”

As a statistician explained it to me, imagine a supermarket: If you want similar products to be near each other, you have a limited number of possible arrangements. In one dimension, products could be side-by-side: diapers with bottles on one side and baby wipes on the other. In two dimensions, related products can also go above and below: on top of the diapers, baby oil; below, diaper rash ointment. But soon, there is no more space and you have to compromise: counting the shelf across the aisle as a third dimension gives room for pacifiers and swaddling blankets, but teething rings and onesies would need to sit farther away. With each new dimension, there are more ways for items to be adjacent—with enough dimensions (precisely: the number of items minus one), everything could be next to everything else, in one dimension or another. This means that, as dimensions proliferate, three-dimensional intuitions about distance do not hold, and many algorithms for calculating distance run into trouble. In high-dimensional spaces, points have more ways to be close to each other. This is the curse of dimensionality: as axes of difference multiply, so do potential connections.[37]

For my interlocutors, this meant that the classificatory promise of big data was also a problem: though tracks could in principle be connected in countless ways, careful pruning was necessary to ensure that these connections made sense. One founder of a music data startup described to me a common problem using audio data to calculate similarity: most audio similarity algorithms will make startling cross-genre connections that humans might not appreciate or be able to notice, seizing on an incidental bit of a piece of electronic music and likening it to a piece of jazz, for example. This posed a problem as, although the music sounded alike "objectively" (i.e. to the algorithm), humans would reject it. To get around this problem required narrowing down the tracks to be compared first. Such algorithms might be fed data that had already been clustered

[37] Elsewhere, I have described this as a form of "hyperactive kinship" for data—with enough data and analytic options, anything can be related to anything else (Seaver 2015,

into genres, only using audio similarity to compare songs within the same genre to ensure that the connections they found would stay within bounds that made sense to people. These pruning processes transformed the "nearly infinite" potential, as Edward had put it, into the actualized subset of dimensions that would be used in practice. Analogizing back to the case outlined by Gupta and Ferguson: though there are in theory infinitely many axes of difference along which people could be sorted, in practices shaped by power relations, only some of those axes were made to matter.

If any kind of space could claim the neutrality once afforded to geographical space, it would seem to be mathematical space, shot through with vectors, dotted by data points, and traversed by algorithms. When anthropologists tried to shift among understandings of space, they did it via mathematical intermediaries. Yet, for my interlocutors, these mathematized spaces posed interpretive challenges, requiring work to make sense as "space." Intuitions about distance did not hold, axes of difference and connection multiplied, and every part of them, from data sources to distance metrics, was up for debate and adjustment. Thus the constitution of mathematical space was a sociocultural enterprise — each decision informed by communal ideas about what these spaces were like and how they should be constructed or tended to. Spatialization required intervention, and what interventions seemed possible or best were informed by their makers' cultural worlds.

Analytic spaces—like the ramifying axes of difference argued for by Gupta and Ferguson or the musical world mapped by Edward — are not neutral or objective substrates, but

---

borrowing a term from Helmreich 2003).

are shaped by their creators' theories about what space is like. For Gupta and Ferguson, space is striated by the politics of difference, for Edward, it is a series of maps facilitating travel. These ideas about space shape the decisions made in the production of analytic spaces, pruning back the growth of connections into manageable, recognizable, and inevitably partial models of the world. To understand the shape these spaces take requires looking for the particular understandings of space that guide this pruning. In the balance of this chapter, I describe one common way of understanding the relationship with space—a set of metaphors that I call "pastoral" because of their relationship to pastoralism as described at the end of the previous chapter and because of their bucolic quality.

We're Park Rangers

At a Billboard conference in 2012, Tim Quirk, who was then the head of Google Music, argued that the massive availability of music online was changing the work of cultural intermediaries. Where the work of classic intermediaries like record store clerks, DJs, and A&R guys was to select and promote music,

> [the] explosion of content has created a new, less sexy need. Telling the entire world what it should and shouldn't listen to has become far less important than simply making this overgrown musical jungle navigable. Online music services need bushwhackers carving paths from one starting point to another. We're not gatekeepers. We're not tastemakers. We're park rangers. (Pham 2012)

In my fieldwork with the developers of music recommendation systems, I heard parts of Quirk's argument repeated again and again. Most of my interlocutors disavowed the power attributed to them by critics of algorithmic systems (e.g. Burkart and McCourt

2006)—they were not tastemakers or gatekeepers, cultural intermediaries who played a determining role in the circulation of music, they said. Instead they saw their task as making the overwhelming scale and diversity of musical space manageable and navigable—as Edward argued, providing maps, or as Quirk suggested, clearing paths. Although one might imagine a world of databases and algorithms to be orderly, Quirk and many of the engineers I talked with saw the space of online music as an inherently unruly wilderness that had to be managed.

Quirk's pastoral language resonates throughout the world of music recommendation. Algorithmic radio stations grow from "seeds" (a suitable descendant of "broadcasting," an earlier agricultural metaphor linking the wide spreading of seeds to the one-to-many sending of messages). Ellie, who worked as a quality assurance tester for a music infomediary, described her job to me as being a "data gardener": she was tasked with pruning the outputs of algorithms to ensure they continued to produce the data they were supposed to and fixing metadata, weeding out inaccuracies. Meanwhile, an engineer working on one of the data sources that fed the algorithms whose outputs Ellie tended to described his job as being a "data plumber," providing the garden with water.

These metaphors extend beyond music recommendation to popular descriptions of machine learning more generally. In his popular press book *The Master Algorithm*, computer scientists Pedro Domingos describes machine learning as being like farming, as opposed to traditional programming, which is like factory manufacture. Where traditional programmers write programs, machine learning programmers write programs that themselves write programs (in a set of prescribed ways). Domingos writes:

> In farming, we plant the seeds, make sure they have enough water and nutrients, and reap the grown crops. Why can't technology be more like this? It can, and that's the promise of machine learning. Learning algorithms are the seeds, data is the soil, and the learned programs are the grown plants. The machine learning expert is like a farmer, sowing the seeds, irrigating and fertilizing the soil, and keeping an eye on the health of the crop, but otherwise staying out of the way. (Domingos 2015, 7)[38]

Critics have noted that, among big data practitioners and popularizers, naturalizing metaphors have become popular to describe the data on which algorithms operate: data is a force of nature or natural resource, a flood or tsunami, oil or gold (Lupton 2013; Puschmann and Burgess 2014; Seaver 2015; Watson 2015). The examples here seem to be in line with this understanding, especially Domingos' contention that machine learning programmers "stay out of the way," letting the system do what it will, consequently guaranteeing a form of naturalized objectivity. Puschmann and Burgess argue that these metaphors misrepresent the processes by which data come into existence. Etymology aside, data is not "given," but "created by humans and recorded by machines rather than being discovered and claimed by platform providers or third parties" (Puschmann and Burgess 2014, 1699). Through these naturalizing metaphors, "the givenness of data is analogized through the givenness of natural resources, which can be mined or grown and which can act as a form of capital with no persistent ties to their creator" (Puschmann and Burgess 2014, 1699). Following this argument, critics might respond to the pastoral metaphor as yet another attempt to naturalize the constructed dataspaces across which algorithms travel.

However, an anthropologically informed resistance to attempts to divide phenomena into "natural" and "cultural" bins suggests an alternative interpretation. Understood in the context of my interlocutors' daily lives, which saw them managing tools, inputs, and outputs produced by others, over which they had limited control, but which they tried to steer toward outcomes they desired, the pastoral metaphor suggests a more ambivalent relationship to nature and control. If Domingos' description of machine learning suggests an idealized life on the farm, where everything works according to plan and farmers simply reap nature's bounty, which is both plentiful and objective, the everyday work of coding bears more similarity to a vision of farming where crops can fail, resource flows can dry up, and growth and balance are the result of concerted, ongoing effort.

Parks, landscapes, and gardens are, in critical ways, *not* natural. To take these accounts of coding as naturalizing errors actually supposes two mistakes: the first is to take data or algorithmic processing as natural, when they are in fact cultural; the second is to assume that because something is natural, it is untroubled. I argue for a reading of the pastoral metaphor not as an ill-informed naturalization, but rather an acknowledgement

[38] This account of machine learning bears certain similarities to more marginal and experimental work in "artificial life," which considers programs to be like organisms

of developers' active role in the production of musical space. The pastoral metaphor locates this work not in an idealized nature, but rather at the intersection of the natural, the cultural, and the technical. Taken as a way of thinking about action and control in this ambivalent middle space, the pastoral metaphor is not simply an error to be corrected — it is a tool that can be used to draw attention to how relationships among persons, algorithms, and data flows are managed and understood. Where the naturalizing metaphors of big data called out by critics misrepresented constructed objects as discovered ones, the pastoral metaphor refers not to objects but relationships between people and objects.

In other words, returning to the broader theme of this dissertation, the pastoral metaphor is an emic way of talking and thinking about algorithmic systems. Neither data nor algorithms nor programmers nor corporations do anything in isolation or complete control. The pastoral metaphor allows for a feeling of control at a remove—not the production of objects but the production of situations that will hopefully bring about certain kinds of objects. While the operation of algorithmic systems plays out over time, it is guided by the choices of the humans within it (themselves patterned in cultural form), but those choices interact with flows of data and bits of computational infrastructure not under their control: surprises still grow out of algorithmic systems, as error or serendipity. These elements of the system are caught up in broader ecologies of meaning and resource flows: a text-scraping bot that gathers key terms from a website breaks when the website is redesigned; a fan club begins streaming songs on repeat, changing the significance of a "play"; a new technique for analyzing audio data is

(Helmreich 1998).

published at an academic conference; another team implements its bit of software in the wrong place and the outputs stop making sense. The data gardeners and park rangers are work not only to provide order according to their particular understanding of order, but to be responsive to these unexpected events that never stop.

If the musical maps I have returned to through this chapter were just maps, then one might think of these problems as diversions that pull the map away from the territory: the territory is still there, not changed by an accident of map-making, and the work of the map-makers is to bring the map back into accord with it. But these "maps," to which developers relate as though they were gardens or parks or farms, are not just maps. They are maps stitched into a territory, a space for action and understanding, where control is neither total nor absent. In spite of the critiques of naturalizing metaphors, which hold that these people pretend that their constructions are like nature, when they are in fact not, these ecologies of control and freedom, maps and territories, bear many similarities to nature. That is, if we understand "nature" not as an idealized, objective, and untouched alterity, but as an outcome of a variety of human and non-human projects.

Nature and Control

In 1994, decades after Volney Stefflre had left UC Irvine, the environmental historian William Cronon organized a seminar there on the topic of "Reinventing Nature," bringing together an interdisciplinary group of humanists and social scientists to investigate the production of "nature." In his introduction to the book that came out of the seminar, *Uncommon Ground*, Cronon spends a long time reflecting on his experience living in a meticulously planned suburbia, where every appearance of “nature” has been hand-crafted. In closing, he turns his attention to the center of campus:

> The campus of the University of California at Irvine is built around a great circular green space called Aldrich Park. Like so many other features of Irvine, it is a carefully planned and constructed place. [...] The paths in the park have been carefully laid out to prevent people from traveling straight across it. They do so quite cleverly, inviting the walker in by means of a well-crafted optical illusion that makes it look as if they do go straight across; only after one is already committed to one’s route is one permitted to see that the lines that at first seemed straight are curved and broken. [...] I have to confess that I found these deceptive pathways rather irritating [...] I could not help seeing these paths as just one more example of the planners’ ubiquitous efforts to control and manipulate my experience of their world, forcing me to conform to their sense of the proper way to appreciate this natural area they had constructed on my behalf. (Cronon 1996, 52–4)

The premise of Cronon’s seminar and *Uncommon Ground* had been that “‘nature’ is not nearly so natural as it seems. Instead it is a profoundly human construction” (Cronon 1996, 25). Nonetheless, certain constructions, like those of the Aldrich Park planners, seemed to chafe more than others. For Cronon, the paths in the park represented attempts by planners to control him, not through any explicit force, but by subtly manipulating his experience of the world.

Cronon’s complaint about Aldrich Park resonates (surprisingly, given its critical tone) with Tim Quirk’s self-characterization of the work of park rangers:

> Being a park ranger means our job isn't to tell visitors what's great and why. Our job is to get them from any given thing they like to a variety of other things they might. We may have our own favorite paths and being park rangers we probably even prefer the less crowded ones, but our job is to keep them all maintained so visitors to our park can chose their own adventure. *They might not feel our hand on their backs as they wander, but it's there. It's just subtle.*

One could hardly ask for a better example of Deleuze's "society of control" (Deleuze 1992) than this: the enclosures of the old music industry, manned by gatekeepers, give way to the apparent openness and freedom of the park, but park rangers subtly keep you on their paths. Deleuze's postscript on control societies—a short essay written late in his life—has been influential for theorizations of how power works online, where the apparent freedom afforded to individuals makes it seem that power is absent (Galloway 2004; Franklin 2015). John Cheney-Lippold ties Deleuze's characterization of the control society to the flexible algorithmic classification techniques that power online targeted advertising, which he calls a "soft biopower": "Enclosure offers the idea of walls, of barriers to databases and surveillance technologies. Openness describes a freedom to action that at the same time is also vulnerable to surveillance and manipulation" (Cheney-Lippold 2011, 177). If hard biopower works through the force of categorization, according to Cheney-Lippold, then soft biopower works at a remove, as "a guiding mechanism that opens and closes particular conditions of possibility that users can encounter" (Cheney-Lippold 2011, 175). It is power not through domination, but through modulation.

It is surprising to hear Quirk actively claim this role for himself, using a metaphor—the hand subtly on the back—which elicits disgusted groans whenever I tell people about it. Typically, this kind of control is considered to be masked by apparent freedom and the denial that any control is operating at all. This is the popular rhetoric of internet platforms, which often claim neutrality while exercising control over the content provided through them (see Gillespie 2010). Yet in Quirk's analogy and the pastoral metaphor more broadly, we find an apparently ambivalent relationship to control on the part of people working within these algorithmic systems: they do not deny their power to shape user experience, to impose their own order on data, or to rearrange computational infrastructures. Rather, they work in a muddier area where maps are sown into territories and they understand their control not as a novel development—as in the historical progression outlined by Deleuze (see the table in Galloway 2004, 114–5)—but as of a kind with other practices as old as agriculture. Their experience of control is not as a master planner from above, but as an interactor within, where their attempts at ordering butt up against recalcitrant others.

One may argue that the current state of online music is better described by Cronon's image of the hyper-manufactured suburban park than by the "overgrown musical jungle" Tim Quirk proposed blazing trails through. After all, the developers of recommender systems for large music streaming services, like Volney Stefflre in his supermarket laboratory, construct environments for their users, control variables, and closely track responses. Their work on musical space is not simply the revealing of latent connections discovered algorithmically. Rather, these hucksters of the symbol also take the cultural order as something to be tended to, broken through, and reconfigured.

These actions are informed by theories about the nature of cultural “space”—architecture, field, or jungle—and the level of control they exert means that their organizing concept-metaphors can become very influential.

Yet, the control (and experience of control) of these people within the algorithmic system is bounded. Counter to critiques that suggest such algorithmic systems are static frames that only provide the illusion of freedom within (e.g. Galloway 2004, who describes collaborative filtering as a “synchronic logic injected into a social relation,”115), these systems actually come to work through interactive, temporally extended processes. While the idealized, unimplemented collaborative filter may rely on an idealized, static understanding of taste, the work of algorithmic systems unfolds over time and in a variety of broader ecologies through which data, people, and techniques move. The pastoral metaphor usefully indexes the kinds of bounded control that are achieved within these systems. Algorithmic systems manifest connections that exceed and precede them, resisting and surprising their human minders: data come already contoured by cultural worlds and the spaces data gardeners tend are not entirely theirs to shape. If we want our critiques to more adequately engage actual practices and to land with the people involved in performing them, we need to recognize the partiality and situatedness of these various forms of control.

CHAPTER 4:
HEARING AND COUNTING

Music Informatic

It is October 2012, and I am standing in the cloister of a Portuguese monastery, listening to a graduate student from New York. We're in the coffee break at ISMIR, the International Symposium on Music Information Retrieval—an annual conference that brings together computer scientists, librarians, and musicologists who study music as a kind of *information*. Although few of the people I talk to consider ISMIR to be their primary conference—they present their main work at conferences on signal processing, musicology, cognitive science, or machine learning—and many worry about its standing in the eyes of their home disciplines, a recognizable core set of participating people and institutions has emerged. The conference is 13 years old this year, and it has hit a groove in terms of the topics it covers: There are a few "symbolic" papers, usually from musicologists, concerned with music as it is represented in scores. There are a few papers from researchers who focus on non-Western music, such as Carnatic music from southern India or Turkish Makam. There are sessions on musical emotion or mood, song structure, small-scale qualitative listener studies, and the analysis of large playlist datasets. One of the most prevalent topics at ISMIR has become what some call computer audition: the science of training computers to hear.

In the computer, an audio file is a long list of numbers for telling speakers how to vibrate—"amplitude and time, that's all audio is," one student tells me. The fundamental task of computer audition is to reduce that series of numbers, typically 44,100 of them a second in CDs or MP3s, to a much smaller set that meaningfully represents the content. These smaller sets are called “feature representations” (in that they represent not the audio itself, but relevant features of it), and they serve as the input for higher-level algorithms, their small size making computation more tractable. Much of the work I saw presented at ISMIR the three times I attended it was focused on developing new representations or testing their performance in a set of benchmark tests: rating the acoustic similarity of songs, automatically identifying genre, or recognizing musicological features like key and instrumentation, for example.

The graduate student from New York—let’s call him Nate—is telling me about his own work in computer audition, which is focused on automatically producing features. After a summer internship at a big software company, Nate has become enthusiastic about a method called “deep learning,” which uses processor-intensive neural networks (so-called because they are loosely modeled on a theory of neuronal structure in the brain) to identify patterns in complex data. Where the current standards for processing musical sound work by applying a series of carefully engineered transformations to audio data to produce purpose-agnostic representations of sound (Nate and others referred to them, dismissively, as “heuristic” or “hand-crafted”), neural networks work directly from the “raw” audio data, trying to derive the ideal feature representation for a given task automatically. This requires a lot of data and computational power, but with the growth of both of those things, neural networks have recently re-emerged as a hot topic in computer science for a variety of applications, and Nate is confident that, soon, work on these resource-intensive methods will pay off. After all, he tells me, “Music is a signal like anything else.”

In this chapter, I examine this understanding of listening. Where previous chapters examined the ways that listeners and music are imagined and interpellated as entities in the algorithmic systems of music recommenders, this chapter investigates a narrower element of those systems: how music "itself" (i.e. its sound) is made manageable for use as data in algorithmic systems. This understanding of sound is particular, and although it is at odds with understandings of sound that have become popular in anthropology, I suggest that there is more to it than a simple quantitative reduction. Thinking of music as "a signal like anything else" leads to interesting consequences for thinking about listening and signals.

To treat music as a signal means to treat it as a series of numbers laden with pattern: "It's all just frequencies at different time scales," Nate says. At one scale are pitches—an A vibrating at 440 Hz. At another, tempo—120 beats per minute. Melody, rhythm, and meter repeat on their own timescales, and at a higher scale is song structure: verse and chorus repeating a few times over the course of a few minutes. One might imagine this progression extending beyond an individual piece of music, with genres and individual tastes representing patterns on an even larger scale. These patterns are latent in the numbers of the data stream, just waiting to be mathematically recognized.

This attitude toward music—that it is essentially informatic, numerical—is not uncommon among the scientists and engineers I conducted my fieldwork with. But, contrary to what one might expect, such an attitude does not lead to a disenchantment or complete rationalization of sound and music. Rather, the equation of music and math has unexpected consequences for how the people at ISMIR and my other field sites interact with, describe, and come to know both sound and the numbers produced to represent it.

The monastery courtyard is covered by a white metal roof and lighting rig, the ground a vast expanse of parquet floor, converted into a conference center. It looks like a high culture backdrop you might see on European arts television. I struggle to hear Nate over the din of the coffee break, as dozens of voices ricochet off the stone walls, metal roof, and lacquered floor before reaching my eardrums. In computer audition, this is called the "cocktail party problem": how do you separate the voice you want to listen to from all the others? The human ear is remarkably good at this, picking out individual sound sources from complexly sounded environments, but computers, like ears as they age, struggle to distinguish signal from noise. I think of how different the courtyard must have sounded when this was still a Benedictine monastery: hushed whispers along the arcades instead of the lively chatter of an annual meeting echoing across a vast parquet floor. Our conversation fades out as they tend to at conferences, dissipating back into the crowd of people looking for old friends and new coffee.

Sound and Sensibility

When anthropologists talk about sound, they are usually talking about the presence of the body. Take, for example, Steven Feld, who has explored the possibilities of an anthropology of sound since his 1982 *Sound and Sentiment*, in work on the poetics of Kaluli song in Bosavi, Papua New Guinea: "Sound, hearing, and voice mark a special bodily nexus for sensation and emotion because of their coordination of brain, nervous system, head, ear, chest, muscles, respiration, and breathing. […] Hearing and voicing link the felt sensations of sound and balance to those of physical and emotional presence" (1996, 97). Feld argues for the potential of "acoustemology," or acoustic epistemology, "of acoustic knowing, of sounding as a condition of and for knowing" (Feld 1996, 97). Paul Stoller writes of the experience of being "penetrated" by sound during his fieldwork on Songhay possession and sorcery (Stoller 1987). Building on Feld's work, David Samuels, Louise Meintjes, Ana Maria Ochoa, and Thomas Porcello have argued that a "sounded anthropology" should pay more attention to soundscapes—sonic environments, analogous to "landscape" in their constructedness, politics, and naturalcultural hybridity. The soundscape concept, they write, "may find more traction in the anthropological mainstream now," with "the return to the body, the senses, and embodiment as areas of anthropological research and sources of local knowledge" (331). Reporting from the Juan de Fuca ridge, 7,000 feet underwater off the coast of Washington, Stefan Helmreich listens to his watery environment, transduced through the instruments and metal walls of a submarine; he uses this claustrophobic soundscape to theorize "transduction," or "the transfer of signals across media," as a mode of "immersive cyborg presence" (Helmreich 2009, 27). Though Helmreich offers transduction "against immersion" (Helmreich 2010), listening transductively remains an engagement with the techno-aquatic surround.

Anthropological treatments of sound are comfortable with what Derrida critiqued as a "metaphysics of presence" (Derrida 1974)—the privileging of knowledge produced by *being there*—which might be expected from a discipline so entwined with a method defined almost entirely by the act of being among the phenomena you are talking about.

In the case of Feld, this is an explicit and extensive commitment to phenomenology; for Samuels et al. and Helmreich, this is no simple faith in the power of presence, but rather a willingness to engage with the politics and techniques that produce sonic experiences of presence—be those struggles over the production of "nature" recordings or the rangefinding operations of submarine instruments.

But what does an anthropology of sound centered around the idea of sound-as-presence have to say to someone like Nate, who spends his days designing algorithms to parse music, switching easily between talk about computers, ears, numbers, and sound? The bodily presence that so dominates anthropological interest in sound is less important for Nate and his colleagues who understand sound instead as a kind of information.[39] "For me," Nate would tell me two years later in a New York bar, "sound exists when it's digitized."

One anthropological script for dealing with such a situation would be to insist on the body—to take Nate by the throat and remind him that it (his throat) is more than a sound source and filter to be modeled in software—it is connected in fleshy ways to the rest of his eating, drinking and breathing life, linked through culture and language to the throats of his colleagues and friends, implicated in circuits of food and air that ground sound in politico-material context. His ears, likewise, are embodied in ways that the history of acoustic science—with its propensity for abstracting or cutting ears out of heads and attaching them to technical devices (Sterne 2003, 32)—fails to appreciate.

[39] Like Feld, however, some of Nate's ISMIR colleagues did take an interest in the significance of birdsong (Stowell and Plumbley 2013).

Another script would be to seize on the contingencies of sound's informatization: information is not found, it is made, and the conversion of sonic experience to numerical data is no neutral process. "Raw data is an oxymoron," as Geof Bowker has argued (Gitelman 2013), and the anthropologist's role is thus to surface the buried contradictions on which informatic edifices are built: the work required to make data appear "raw," the practice and mess left just out of formalistic frames. Nate's narrow attention to digitized sound misses the necessary embodiments and excesses of sounds that move through space and time, reducing them to quantities that almost definitionally cannot capture those qualities of sonic experience anthropologists find important.

Over the course of my fieldwork with scientists and engineers like Nate, these anthropological scripts informed my interactions: I would raise questions about contexts I saw elided from numerical discourses, point out the shortcomings of formal representations, and generally advocate for context and presence, making myself a persistent but polite pain in the ass. Strangely enough, my interlocutors rarely disagreed with me. They acknowledged the limits of their representations, the overly broad claims based on overly narrow data, and the scope of sonic worlds that exceeded their data practices. Then, they turned around and went on with business as usual. This was a challenge for critiques premised on the idea that they revealed fundamental flaws in these practices—what was I to do when these arguments didn't seem to have effects?

My response was a classically anthropological one, inspired by the ethic of interpretive charity and the injunction of methodological philistinism outlined in chapter 1. My goal was not simply to remind my interlocutors of the shortcomings of their worldview, but to try and hear how notes that seemed discordant to me fit together for them—to learn to listen like they did, to aim for a sonic version of "the native's point of view" (Geertz 1974).This was a kind of resonant anthropology which did not proceed from the premise that sound is essentially about presence, bodies, and environment (see, e.g. Samuels et al. 2010). Why, I wondered, should I privilege my own anthropological understanding of sound over the informatic understanding of my interlocutors? Or, to put it another way, what did the assumption that I knew what sound *really* was stop me from hearing in the field?

We have already heard from Nate at ISMIR and from the nascent literature in sounded anthropology. Next, I turn to historical precedents for linking hearing and counting, which resonate with sociality as well, in the work of physicist Hermann von Helmholtz and sociologist Gabriel Tarde. Then, I return to ISMIR and the work of making musical feature representations, paying close attention to the production of one of the most popular feature representations, called "mel-frequency cepstral coefficients." At a conference session after my conversation with Nate, the presenters offer new representation techniques, which they demonstrate using sonic illustrations, reversing the typical understanding of quantification as a reduction from the fullness of actual sound. I argue that this unusual acoustemology (knowing through listening) can be understood by thinking of the relationship between music and its quantification as a kind of resonance, which goes both ways: the connection between numbers and sound does not leave numbers unaffected. I conclude by turning to the "deep learning" Nate advocated for in the monastery courtyard—a style of quantitative hearing that grew in popularity during my fieldwork, which attempted to parse "raw" audio data directly, often trained by the "cultural" data of listening behavior. To make sense of the outputs of these notoriously opaque systems, my interlocutors (and I) learned to listen to algorithms, trying to tune our attentions to each other's and the computer's.

Math and Music

> *I am devoted to dispelling the widespread myth that Greek mathematics was developed for its own sake. For nothing could be further from the truth. Mathematics was invented for music.* -Friedrich Kittler (2006)

The identity between hearing and counting that Nate used to make sense of his work has a long history, particularly in the study of harmonics, which appear in both mathematical and musical contexts dating back to Pythagoras, and were elaborated in the early seventeenth century by the French theologian and mathematician Marin Mersenne. As the historian of science Olivier Darrigol has argued, "The occurrence of the same word ['harmonic'] in musical and mathematical contexts is neither a coincidence nor a purely metaphorical effect" (Darrigol 2007, 343). Rather, "acoustic theories for the emission, perception, and propagation of sound constantly bridged musical and mathematical harmonics" (Darrigol 2007, 344). Mathematics and music did not find each other recently, but rather informed each other's development over millennia (Crombie 1990, 363–378). With the development of modern acoustics and otology (the study of the ear), which happened alongside the development of new communications technologies for transmitting sound like the telephone and the phonograph (see Sterne 2003 and Mills 2010), this mathematical understanding subtended an emergent science of hearing. Over the development of European acoustics, hearing came to be understood as a peculiarly mathematical sense and mathematics a peculiarly auditory science. One of the foundational sites for the modern science of hearing was nineteenth century Germany (Jackson 2006), where the German polymath Hermann von Helmholtz elaborated an influential model of hearing.

In the winter of 1857 Helmholtz presented a lecture in Bonn, on "The Physiological Causes of Harmony in Music." In his lecture, which touched on the connections between music, physics, and the anatomy of the ear, he gave an unusual description of a concert:

> From the mouths of the male singers proceed waves of six to twelve feet in length; from the lips of the female singers dart shorter waves, from eighteen to thirty-six inches long. The rustling of silken skirts excites little curls in the air, each instrument in the orchestra emits its peculiar waves, and all these systems expand spherically from their respective centers, dart through one another, are reflected from the walls of the room and thus rush backwards and forwards, until they succumb to the greater force of newly generated tones. (Helmholtz 1995, 57–58)

The world is awash in vibration and resonance; it is "a variegated crowd of intersecting wave systems" (Helmholtz 1995, 57). Hearing, according to Helmholtz, was the privileged sense for disaggregating the crowd into its constituent parts:

> Although this spectacle is veiled from the material eye, we have another bodily organ, the ear, specially adapted to reveal it to us. This analyzes the

> interdigitation of the waves, [...] separates the several tones which compose it, and distinguishes the voices of men and women—even of individuals—the peculiar qualities of tone given out by each instrument, the rustling of the dresses, the footfalls of the walkers, and so on. (Helmholtz 1995, 58)

Helmholtz would become famous in the field of acoustics for his development of this idea—the "resonance model" of hearing. The hypothesis was that the ear performed a physical version of Fourier analysis, a mathematical method for decomposing a complex wave into mathematically simple ones. What Joseph Fourier had proven years earlier was that any waveform could be represented as the sum of a (potentially infinite) series of simple sine waves. Helmholtz argued that hearing functioned in the same way: a complex waveform reached the eardrum and the inner ear, which contained a set of elements resonant at different frequencies, physically separated the wave into its parts.

If Helmholtz blurred the line between sensation and mathematics, he also entangled the decomposition of complex signals with the identification of social distinctions. In his evocative concert scene, the ear not only decomposes sound into its basic frequencies, it also distinguishes among instruments, male and female voices, and musical and non-musical sound. The ability to distinguish among tones and the ability to distinguish among social categories—gender and noise—are linked for Helmholtz in the biomechanics of the ear, and he hints at the possibility that the resonance between math, music, ears, and sound might extend even beyond them, to the social scene of sonic action. Hearing is simultaneously biological, numerical, and sociocultural: it resonates with stiff hairs in the ear, the mathematics of sine waves, and the vibrating entities that populate the world.

Where Helmholtz hints at the sociocultural consequences of the resonance model, an elaborated version can be found in the sociology of the statistician Gabriel Tarde, the sometime rival of Émile Durkheim, who proposed a vibrational ontology bridging sensation, sociality, and statistical description. Writing in his Economic Psychology in 1902, Tarde describes a vibratory world that resonates with worlds I've described so far:

> Everywhere there are harmonies which repeat themselves: a wave is actually a harmonious succession of movements, equilibrium in motion, falling back on itself like a musical phrase. (Tarde 1969, 143)

Although Tarde's preoccupation was with social processes of imitation, he understood these processes in the broader context of what he called "universal repetition." In the introduction to the English translation of Laws of Imitation, Franklin Giddings described Tarde's interest like this:

> M. Tarde perceived that imitation, as a social form, is only one mode of a universal activity, of that endless repetition, throughout nature, which in the physical realm we know as the undulations of ether, the vibrations of material bodies, the swing of the planets in their orbits, the alternations of light and darkness, and of seasons, the succession of life and death. Here, then, was not only a fundamental truth of social science, but also a first principle of cosmic philosophy. (Giddings 1903, v)

Repetition and vibration were central to Tarde's social theory, and they put the social world in direct contiguity with the natural world. For Tarde, the world is all vibration at different scales—oscillations of light in the ether, animal populations on the savannah, sound in the air, crime rates in Paris. These repetitions were fundamental to Tarde's understanding of quantification and science. For Tarde, a world of vibration was also a world of quantities, repetitive and thus measurable. Without repetition, he argued, there could be no quantification—no accumulation of like units to be compared—and thus no science. Tarde justified his sociology by arguing that social processes shared repetitive features with the objects of other sciences.

The contiguity Tarde saw between the social and natural sciences is perhaps most evident in his comparison of statistics and the senses. "Let us take any graphical curve," he wrote in "Archaeology and Statistics,"

> that, for example, of criminal recidivists for the last fifty years.... Is it not like the sinuous lines, the sharp rises and sudden falls in the flight of a swallow? Why should the statistical diagrams that are gradually traced out on this paper from accumulations of successive crimes and misdemeanours ... be the only ones to be taken as symbolical, whereas the line traced on my retina by the flight of a swallow is deemed an inherent reality? (Tarde 1903, 132–3)

Here Tarde hints at the semiotic consequences of his vibrational monism: there is no fundamental distinction between the symbolic motion of statistical figures and the indexical motion of light on the retina. Their difference is not of type but of degree. The indexical is no more real or less arbitrary than the symbolic, only faster. Statistics were laborious to interpret and delayed from the phenomena they described; eventually, Tarde supposed, statistics would continue "to gain in accuracy, in despatch, in bulk, and in regularity" (Tarde 1903, 133) until this difficulty was overcome, and "a statistical bureau might be compared to an eye or ear" (Tarde 1903, 134).[40] And while Tarde figured statistics as sensate, he also figured the sensorium as statistical, extending the quantitative understanding of hearing we have heard so far to the other senses:

> Each of our senses gives us, in its own way and from its own special point of view, the statistics of the external world. … Every sensation—colour, sound, taste, etc.—is only a number, a collection of innumerable like units of vibrations. (Tarde 1903, 34–5)

Bruno Latour describes Tarde's argument as "a progressive fusion between the technologies of statistical instruments and the very physiology of perception" (Latour 2010, 156).

Taken together, Tarde and Helmholtz suggest an overarching resonant sensory epistemology that, while elaborated in the context of the ear, extends well beyond it. For Tarde and Helmholtz, as for Nate, counting is a fundamental fact of perception, as is resonance—a sensitivity to pattern. Where Helmholtz offers a resonant connection between hearing and counting, Tardean statistics gives a model for thinking about the knowledge of sound and the social together, locating sonic knowledge practices in a social context that is not external to questions of number, vibration, or resonance, but rather deeply implicated in them.

> "Mathematics and music!", Helmholtz said. "The most glaring possible opposites of human thought! and yet connected, mutually sustained! It is as if they would

[40] Tarde's prediction bears more than a passing resemblance to the popular "3 Vs" definition of "big data" produced by technology consultancy Gartner: "velocity, variety, and volume." I draw the connection between these epistemological practices and big data more explicitly in the conclusion.

demonstrate the hidden consensus of all the actions of our mind" (Helmholtz 1995, 46–47)

Mel-frequency Cepstral Coefficients

Back in the monastery, I am sitting in a conference session on "Audio Classification." The four papers being presented in this session — from research labs in San Diego, Montreal, Barcelona, and Stanford—offer different techniques for organizing audio files based on their content. These techniques rely on feature representations to turn raw audio data into meaningful and manageable smaller sets of numbers, on which classification algorithms can work. As Tarde said of statistics, feature representations "save us trouble by synthesising collections of scattered homogeneous units for us," giving us "the clear, precise, and smooth result of this elaboration" (Tarde 1903, 134). That is to say, they identify patterns in sound data and represent those patterns with a smaller set of numbers. Feature representations serve two purposes: they simplify data, making computation more feasible, and they reduce "raw" audio data according to models of human hearing, better reflecting how musical signals are eventually perceived. They are quantitative summaries of sonic patterns.

The standard feature representations used at ISMIR are called "mel-frequency cepstral coefficients," or MFCCs. Borrowed from the speech recognition research community, MFCCs are "short-term spectral based features" (Logan 2000, 1), meaning that they describe short segments of sound in terms of the audio frequencies that make them up. As Beth Logan wrote in an early ISMIR paper assessing MFCCs' utility for music analysis, "Each step in the process of creating MFCC features is motivated by perceptual or computational considerations" (Logan 2000, 1)—that is to say, the MFCC is a feature representation which adjusts both to the peculiarities of human hearing in the ear and brain (or "psychoacoustics") and to the particular efficiencies of computers.

One of the weaknesses in Tarde's vision for the sensory future of statistics is that he glosses over the struggles that always attend the production of numbers. The conjunction of counting and hearing is by no means effortless or uncontentious. The sound of a piece of music may seem "objective" relative to subjective matters like someone's taste for it, perhaps providing a stable ground on which to build an algorithmic classifier. However, disputes over how to parse audio give lie to the idea that sound is a strictly objective signal to be incorporated into algorithmic systems (as opposed to, say, listener behaviors that must be interpreted as indicators of subjective taste). The decision to use sound as an algorithmic input requires choosing among a variety of possible representations, and the criteria by which this choice is made are a matter of significant public debate in the research community.

The widespread usage of MFCCs in the ISMIR community has made them a target for scrutiny, and three of the papers in this panel on classification offer alternative feature representations meant to capture aspects of the audio signal that they claim MFCCs neglect. Because of the widespread use of MFCCs (and their standing as the typical representation against which new ones are evaluated), it is worth examining in detail how they are made. So, below, a tour of the numerical steps that go into making a typical MFCC representation, adapted and elaborated from (Logan 2000).

1. Take a "raw" audio signal.

The basic assumption of digital audio is that any sound can be represented as a sequence of numbers that describe the amplitude of a wave over time. The sample rate dictates how many numbers per second are used: 44,100 is typical for CDs and MP3s. Each of those numbers can be a limited number of values: in the case of CDs that number, known as the bit depth, is 65,536. So, one typical second of sound will be represented by a sequence of 44,100 numbers that range between 0 and 65,536.[41]

2. Divide the signal into short, 20 millisecond windows.

Representations of sound as a waveform are said to be in the "time domain": time goes from left to right, while the line of the wave traces out changes in intensity. Researchers are typically interested in the component frequencies that make up a given sound, so they want to convert this signal from a line in the time domain to a spectrogram in the "frequency domain"—a kind of representation that shows which frequencies are present at a given moment and their relative intensities. To convert to the frequency domain, the signal is first sliced into small frames to be analyzed. Now, we have a set of 20 millisecond-long snippets of audio consisting of 882 numbers each. (This is a slight simplification, as it turns out to be useful to have our windows with blurry, overlapping edges, but it suffices for my purposes here.)

[41] In ideal circumstances, this kind of high-resolution audio is the input to the MFCC process, but in some cases, the source audio has already been altered in some way, e.g. by MP3 compression. Recent work presented at ISMIR has examined what effect these alterations have on the usefulness of MFCCs (Urbano et al. 2014).

This windowing—like any other step in the production of numbers—is occasionally a cause of dispute within the academic community: windows that are too long may blur distinct sonic events together, and windows that are too short may miss lower-frequency waves. At a previous ISMIR meeting, I watched a presentation about Indian classical music, which is said to include many microtonal intervals. The researcher suggested that his analysis, which examined the frequencies present in small time windows, indicated that these microtonal notes were not present in recordings. An audience member noted that perhaps his windows were too large: these notes were often supposed to be used as grace notes, played very briefly before other notes, and if a window contained both the microtonal grace note and the note it led into, the former would be drowned out. Another dispute centered on the windowing practices used by The Echo Nest: rather than breaking the audio up into an even grid, their windows were derived from an "event detector," which aimed to identify changes in the musical audio signal. Thus, in theory, window sizes would change dynamically to hold single musical events. However, as academic critics pointed out at the conference, this strategy was vulnerable to errors in event detection, and because Echo Nest analyses were proprietary, such errors could not be readily identified by outsiders.

3. Apply a discrete Fourier transformation to each frame.

Now we are ready to transform our 882 numbers from the time domain into the frequency domain, by applying a bit of math known as the discrete Fourier transform. This decomposes our single waveform into 256 frequency bins (the discreteness of these bins is why it is called a "discrete" transform), vibrating at various intensities—a vertical stack of numbers representing the strength with which frequencies were present in the original signal. The math of the Fourier transform is more than we need to get into here, but an example: if our original sound was a mixture of three pure tones — an A, with a G below it and an F sharp above it, say—the time domain representation would be a complex wave, moving up and down as the tones interfered with each other. In the frequency domain representation, however, the underlying simplicity would be revealed: we would see bright spots at the bins containing 440 (A), 196 (G), and 740 (F sharp), reflecting the fact that our original signal was composed of waves at those frequencies. If we were to put all our 20 millisecond slices next to each other, we could see those pitches move over time, from chord to chord, for example, in an image called a spectrogram. Spectrograms can show underlying patterns that are not obvious from the waveform, but which are readily interpretable, at least in simple cases. Now, our frame is just 256 numbers that describe the frequency content of 20 milliseconds of sound.

Although the waveform, which represents sound in the time domain, is a popular way to represent sound among the general public (Walker 2011), spectrograms are the typical representation found among music informatics researchers. One university lab provided its students with t-shirts that had a 3D spectrogram image printed on them along with the jokey slogan: “The time domain is for losers.”

4. Adjust the intensities in the frame logarithmically.

The 256 numbers in our frame represent how strongly frequencies are present in the signal. However, human hearing does not register this strength directly as loudness. Instead, slight differences at the bottom of the scale make more of a difference than those at the top. “The perceived loudness of a signal has been found to be approximately logarithmic" (Logan 2000, 2). So, the perceived difference between 5 and 10, say, is much larger than the perceived difference between 100 and 105. This logarithmic transformation makes the numbers more closely resemble how loud they would sound to a human listener.

5. Smooth the 256 frequencies into 40, according to the mel scale.

Now, we want to reduce our 256 numbers down further, to more concisely represent the sound. To do so, we will group them into 40 bins. But, as with signal strength and perceived loudness, human perception of frequencies does not directly track with their numerical values. To account for this, we space out the bins according to the "mel scale"—a scale derived from experiments that show that humans perceive pitch linearly below about 1000Hz and logarithmically above it. This gives us a set of increasingly large bins, and we average together the frequencies that fall in each of them, ending up with 40 numbers, adjusted for human perception of loudness and pitch, representing 20 milliseconds of sound.

These last two steps adjust for human perception, to help a computer hear like a human would, at least in terms of quantitative scale. They rely on psychoacoustic research (the source of the mel scale and the knowledge of human responses to sonic amplitude), which, if it were different, would result in different outputs. Indeed, the mel scale is itself contested as a reflection of human perception (see, e.g., Umesh et al. 1999).

6. Decorrelate the remaining audio features.

As described in step 3, converting from the time domain to the frequency domain is a way to reveal patterns that are not obviously apparent from the waveform—a frequency that pushes the waveform up every .003 seconds is hard to see on a complex waveform, but easy to see on a spectrogram. Researchers are interested in even higher order patterns, though. Most naturally occurring sound sources produce identifiable clusters of frequencies (for instruments, this is called "timbre," and refers to the different set of frequencies produced by, say, a violin and a trumpet playing the same note). Because of this, certain frequencies often occur together in patterned ways, and these patterns are hard to see on a spectrogram. This patterning means that the 40 numbers we have so far are "highly correlated" with each other. Their being correlated means that they could be more efficiently represented by a smaller set of numbers (making computation easier); it also means that there is a deeper underlying regularity in the sound (which humans are considered able to hear).

To "decorrelate" these values, we transform them into the "cepstral" domain, effectively repeating the process used to go from the time domain into the frequency domain (this is also why it is called "cepstral," an anagram of the "spectral" representation we produced in step 3). Where a bright spot in our spectrogram would indicate the presence of a given frequency (i.e. a pattern in time), a bright spot in the cepstrum indicates a pattern in the frequencies. These numbers are, at last, the mel-frequency cepstral coefficients, and each of them statistically "explains" a different aspect of the variance in our previous numbers. We can get many coefficients (up to nearly as many numbers as we had to explain in the first place), but, like typical researchers, we may decide that after 13 numbers, the amount of variance remaining to explain is trivial. So now, we have 13 numbers that describe higher-level patterns of frequencies in 20 milliseconds of sound. These numbers provide the material on which various listening algorithms work.

***

The algorithmic process of transforming "raw" audio (which has already been formatted by a variety of standards for digital sampling) into a small set of numbers amenable to large scale processing and relevant to human perception can be seen as another technique for mediating scalar imbalances between people and their media environments, and it involves much more than code. The steps in the construction of the MFCC combine concerns about human and computer capacities, performing transformations that are contested and drawing on theories that are contested. The "algorithm" seen in this light is not a simple sequence of steps, but rather a sequence of choices, all of which can vary and whose variations are guided by ideas about salience that are themselves variable.

MFCCs are thus an example of Jonathan Sterne's "perceptual technics" as described in chapter 1: technologies that integrate experimental knowledge of human perception with the computational (and often commercial) exigencies of technological communication (Sterne 2012). Though Sterne developed the term to describe the MP3 format, it applies as much or more to MFCCs, which can be understood as an extension of the MP3's logic of compression. In them, human hearing and computational analysis blend and, frequently, are found to require similar concessions. They are extraordinary accumulations of scientific facts, conventions, and understandings about listening. MFCCs push perceptual technics beyond the realm of the audible: they represent salient auditory features in such a condensed form that they cannot be played back as sound—at least not directly.

Resynthesis

Back in the session on “Audio Classification,” the presenters are arguing that MFCCs are poor representations for the tasks to which they have been set. The current presenter suggests that, because MFCCs were developed for speech recognition, they neglect musically salient features like pitch. Another argues that, because MFCCs are usually calculated from such small slices of time, they miss musical structures that develop over longer time-scales. They offer their own alternate representations, designed with particular tasks in mind, and compare their performance to MFCCs. The current presenter is attempting a task that requires the algorithm to correctly tag a large set of songs with labels like “Rock,” “Piano,” “Electronica,” and “Exciting.” He shows how well the classifier performs when using MFCCs versus his own new feature representation — unsurprisingly, his feature representation outperforms the MFCCs. But he does not stop with the quantitative proof. Like the two presenters before him, he demonstrates the differences between his own feature representation and MFCCs through a sonic illustration.

Remember that these feature representations typically do not contain enough information to be played back directly, so in order to illustrate what these sets of numbers sound like, they must be “resynthesized” to be turned back into sound. The skeletal numbers of the MFCC, for example, are fleshed out by adding white noise to the signal. The presenter plays an excerpt of Carole King's "You've Got a Friend" over the speakers. Then, the MFCC resynthesis. The piano notes sound like cavernous drum beats, and King's barely recognizable voice hisses tuneless over them: *Close your eyes and think of me* / *And soon I will be there*. Imagine Carole King, but when she opens her mouth to sing vowels, only a rush of white noise comes out. The sound is terrifying and the audience laughs. When he plays back his own features, they sound much closer to the original — like a low-quality MP3 rather than an ominous message from the cepstral realm.

During the question and answer period, an audience member points out that, technically speaking, it doesn't matter what these representations sound like—it is possible that MFCCs numerically contain information about pitch that is simply not recreated in the white-noise resynthesis process, and the real test of a representation's adequacy is how it performs in computational tasks. What a given representation sounds like when turned back into sound is essentially irrelevant to a computer's ability to make sense of it; this representation may very well suffice for the computer to identify pitch, even if the resynthesis seems evidence to the contrary. Carole King's wicked metamorphosis probably has more to do with the acoustic features of white noise than the adequacy of MFCCs. The presenter agreed with this criticism, but suggested that "the point was to hear what was lost in the transformation"—to offer a sonic example that illustrates (or "resonates with"), rather than conclusively demonstrates, the claim that one feature representation is better than another. This style of demonstration reappeared at the two other meetings of ISMIR I attended, and it does not seem to be going away, in spite of a general agreement that sonifications of feature representations do not conclusively prove anything.

The persistence of these sonic illustrations is thus something of a puzzle. Why do music information scientists persist in making these sonifications that they do not consider scientifically valid? Why do they want to hear the mathematical creations they have engineered?

Quantitative Acoustemology

Practices of listening to computational products kept popping up throughout my fieldwork—interviewees at technology companies in San Francisco told me about how they listened to the playlists generated by software they wrote to see if the system "worked," and they learned to recognize the signature playlisting styles of major algorithmic radio companies; the people who sat next to me during my internship checked the work of classifying algorithms by listening to the songs they had sorted; and at weekend "hackathons," where amateur and professional coders built prototype software, their performance was constantly assessed through headphones (in what were called "sniff tests" or "smoke tests," pointing to the lack of sonic metaphors available for this kind of knowledge work). The tentative testing that characterizes contemporary software development—write code, run code, get error message, adjust accordingly, repeat—bears an interactional structure not unlike musical composition—write, play, listen, write repeat—and new software structures, like new compositional ones, shorten those intervals or do away with them in favor of improvisational "live" modes of creative code/musical expression. Listening, literally and metaphorically, becomes a central element in the feedback cycle of creative production.

This interpretive work—the importance of human perception and judgment within algorithmic procedures—is a key element in cultural influence on algorithmic systems. Before they ever make it to a user-facing recommender system, audio feature representations are tuned to human hearing, through both the quantitative evaluation of algorithms informed by research on hearing and through the literal listening practices of their creators. These listening practices do not stop once the audio data is piped into the recommender system: outputs are listened to and evaluated and systems are changed as a result. The tuning of an algorithmic system is more like the tuning of a musical instrument than one might think, requiring the tuner to listen to outputs and adjust them until they come into line with culturally informed expectations. These expectations can be understood as “tastes,” or more abstractly, as acquired sensibilities that guide interaction with the world—fitting with Antoine Hennion’s understanding of taste in art or Tim Ingold’s account of “skill” in technology, as described in chapter 1.

Returning to the term from Steven Feld, these listening practices are a kind of “acoustemology”—a knowing through hearing. Feld developed his ideas about acoustemology in the rainforest soundscapes of Papua New Guinea, to argue for an embodied, environmental, and multisensory style of knowing—the ability to make sense of sound, to hear predator, prey, and people is crucial in the jungle. However, as I’ve described in this chapter, acoustemological entanglements exist even in the setting of western science, which tends to disregard the embodied, experiential quality of sound that Feld finds so important in favor of numbers. This quantitative acoustemology poses a challenge to common ideas about scientific rationality and to the notion that sonic knowing is somehow opposed to it.

Helmholtz's quantitative ear and Tarde's sensory statistics provide a way to make sense of what is going on here. The link between sensation and quantification is not a one-way street, in which all phenomena are reducible to numbers. Rather, there is a resonant exchange between counting and sensing—remember that for Tarde, not only are the senses statistical, but statistics is sensory, and for Helmholtz, music and mathematics resonate as equals. Understanding the functions of algorithms and ears as analogous, computer audition researchers hold them together. It becomes important to them not only to prove their math, but to hear it. Scientific ideals that dictate the purity of methodically produced numbers conflict with the analogical thinking that linked hearing to computing in the first place. Although numbers proliferate in computer audition, the evaluative capacities of the human ear are still in play. Hearing is something you can count, but counting is also something you can hear.

Quantification can be understood as a style of transduction (Helmreich 2008)—a technosocial interface between numerical and acoustic domains. Like other transducers (see e.g. Sterne 2003 on the phonograph), quantification is frequently simplified, naturalized, and taken as objective. Understandings of transductive technologies and ears resonate with each other. As John Durham Peters has written, "human-machine mimesis is mutual" (Peters 2004, 189)—the casting of the ear as a frequency analyzer and the understanding of frequency analyzers as like ears go hand in hand. Computer audition and human audition influence each other, as though through a kind of sympathetic magic. “Metaphors leap off of pages—and out of ears—into machines” (Peters 2004, 187). However, as my fieldwork with researchers in music informatics has demonstrated, there is not a simple, self-evident conjunction between techniques for counting and hearing. Rather, it is a contested site of translation, where quantitative/transductive options abound. As researchers construct the higher order feature representations of computer audition, they argue among them by drawing on both perceptual and numerical justifications. We might agree that hearing and counting are connected, but we need to acknowledge that there are many ways to count and many ways to hear.

This strangeness—the opening up of both what it means to count and to hear to dispute and inventiveness—resonates with recent anthropological work on number as “inventive frontier” (Guyer et al. 2010). This work contests the common humanistic understanding of numbers as a force for rationalization and disenchantment, drawing on a range of empirical work that demonstrates the inventiveness, strangeness, and irrationality that attends the use and production of numbers in sites around the world. Rather than embracing the divide between the quantitative and the qualitative or the scientific and the humanistic, such an approach draws into question this very foundation, bringing the “destructive analysis of the familiar” Edward Sapir took as the basic heuristic of anthropology to our very methodological framing assumptions (Sapir 1921, 94). What does ethnography sound like if both sound and ethnography are not necessarily opposed to the formalizing, quantifying practices they have come to be understood against?

The Neural Net

At the beginning of this chapter, Nate—the graduate student from New York—was telling me that "hand-crafted" feature representations like the MFCCs and its alternatives described thus far should be replaced by "feature learning" performed by neural networks. These algorithmic systems would do away with the painstaking work of determining which transformations to make on audio data to best reflect human hearing. Where MFCCs are purpose-agnostic—theoretically driven representations of sound that are intended to be used as input for any number of tasks, like segmenting songs into parts or tracking the melody—the representations produced in neural networks are created in response to particular tasks. Given "raw" audio data and an objective, such as correct classification into genre, the system learns a feature representation that optimizes performance at that task, according to some predetermined metric. The features thus developed are tuned specifically to specific tasks, rather than to a more generic idea of hearing, and they are notoriously opaque to the people who build them.

These systems proceed without any explicit theorizing about hearing. Feature learning systems thus seem to escape the entanglements of perceptual technics that mark standard feature representations. Unlike MFCCs, they do not incorporate psychoacoustic research, and they are often built by people who work across domains (a setup used to learn features for music genre classification could later be tweaked to learn how to classify images of plankton or galaxies). Rather, they are elaborate correlation finding machines, learning mappings between different patterns: between the numbers that represent "raw" audio data and those that represent user-generated genre tags, for instance. However, as I found over the years after I first talked to Nate, quantitative acoustemologies remained central to how the "hearing" of neural networks came to be understood.

***

It is July 2014, and I am at a music hackathon at the Spotify offices in New York City, listening to an arcane subgenre of heavy metal. Scattered at tables around the large room, people are working on their weekend projects—sensors that turn dances into songs, synthesizers built inside of web browsers, and at my table, a deep neural network that had apparently learned to hear the differences between musical genres.

I am sitting with a graduate student I know from a previous hackathon—let's call him Tomas. Tomas has long, red hair and a wardrobe consisting mostly of band t-shirts from his favorite genre: "Djent." Djent is a subgenre of metal, named onomatopoetically for its signature guitar sound. Djent, he tells me, is a kind of "technical metal"—rhythmically intricate stuff, defined by its precision and technique. Identifying it (not to mention appreciating it) poses a challenge to computers and non-metal fans alike. To the untrained ear, like mine, it sounds a lot like many other subgenres of metal: there are chugging distorted guitars, the singer shouts lyrics about darkness and reality, it has a combative relationship to religion, and so on.

Tomas is getting a PhD in machine learning, using neural networks to classify music based on audio data. Among the music information scientists I talked with during my fieldwork, music classification based on audio is a notoriously difficult problem. When you ask a computer to classify the subgenres of metal based on their sound, it makes mistakes that no true fan would make. I offer Tomas an anthropological explanation for why this might be: perhaps the differences among subgenres are more cultural than auditory, dependent not on their sound, but on the scene they come from. Indeed, other music informatics researchers had suggested similar explanations to me for why this task was such a challenge: often the distinction between two genres most obviously manifests in the clothes people wear more than the sound, making the division into genres effectively arbitrary with regard to sound, though it may be quite stark in people's listening.

Researchers like Tomas call this problem the "semantic gap"—a divide between the information contained in a song's sound and the meaning humans derive from it. Tomas, however, is insistent: if he could learn to hear the difference between Djent and other styles of metal, a computer should be able to as well. With enough of the right data, it should be theoretically possible to jump over the semantic gap. Conveniently for Tomas, he had spent the summer as an intern at a large music streaming company, working with their machine learning team. As described in chapter 2, these teams are drowning in listening data. They have massive Hadoop databases that store every listening event — every play, pause, and playlist—from every listener. What Tomas has done is to bring this listening data and the audio data together, using a deep neural network, running on a customized set of processors in his office. This system has "learned" a relationship between patterns in audio and patterns in listening behavior, for the million most popular songs on the service.

“Pick a number between 0 and 2047,” he tells me. I pick one and the computer spits out a playlist: coincidentally, it is some kind of metal. Tomas plays a bit from the middle of the first track for me, identifying it as a subgenre called metalcore, and he draws my attention to some of its defining features: the rapid-fire drum pattern known as a blast beat, the style of singing, the guitar tone. Together, we try to pick out the common sounds across the tracks in the list, to imagine what the computer has heard. He points out patterns he has noticed and plays a few more tracks.

The 2048 options Tomas gives me correspond to filters in one layer of the neural network. He thinks this is where the system has managed to learn to discriminate among subgenres. He’s gone through the first 50 or so and tried to identify them, playing some of them back to me: I hear Chinese pop, Christian rock, and even female comedians. The effect is uncanny—it sounds like he computer has learned to listen with a culturally trained ear: not only comedians, but female ones; not only rock, but Christian rock. The similarity seems to extend beyond easily described qualities of the sound into subtle shadings of timbre—there is something similar about the deliveries of the comedians and the ways their audiences laugh; one playlist is composed exclusively of electronic dance music by a single producer—Armin van Buuren, whose signature tone plagiarizes itself across dozens of tracks.

It is an astonishing experience, and later that night I write in my field notes: “I’m wondering if this is the weekend I lose my incredulity: this thing seems like it *works*.” The methodological philistinism that had guided my research in an abstract sense was easy to maintain when systems seemed to fail in obvious ways regularly; this system that seemed to have achieved an incredible performance threw my cynicism off balance. So, I began to consider how my sense of wonder had come about. What processes led me to the feeling that Tomas's neural net actually understood musical genre? My listening exercise with Tomas had been an example of what I would come to think of as "learning to listen."

Learning to Listen

Neural networks like these are a favorite example for an argument that critics of algorithmic systems like to make: sometimes, algorithms are so complex that even their makers don't know how they work. Unlike the clear steps of an MFCC, the choices involved in a neural net's "hearing" are hard to see. They are essentially elaborate counting systems, passing values around between a large number of "nodes" that perform simple transformations many times. Although the nodes of the neural net perform very simple mathematical operations, they do so many of them, and in such dense interaction with other nodes, that extracting a plain-language "X is Y therefore Z" kind of explanation for their output is challenging. Critics have gotten a lot of mileage out of this point: apparently straightforward and rational systems have surprising emergent properties; algorithmic logics exceed human cognition; we need new regulatory paradigms to hold such incomprehensibilities in check. If a computer-science-101 algorithm is easy to see and comprehend, a neural network is an honest-to-god black box, as inscrutable as the interior of someone else's mind (Duranti:2008).[42]

[42] This is perhaps the most defensible parallel between neural nets and actual brains: although their simple operations and elaborate connections are nominally inspired by

But while critics have noted the limits of rationality and barriers to comprehending algorithmic action, they have paid less attention to what people like Tomas do in the face of these interpretive challenges. Faced with challenging complexity, they do not give up and put blind faith in the objectivity of the system, but rather they interpret. In the case of music, they *listen*, and understanding how this listening works is critical for understanding how machine learning systems come to work as they do.

Machine learning researchers have a term for the problem of understanding what complex algorithmic systems spit out: interpretability. An output is interpretable, if, like a playlist made entirely of metalcore, it seems to make sense. An uninterpretable output doesn't make obvious sense: some of Tomas's 2048 playlists, for example, seemed to be composed of multiple genres at once, that had nothing to do with each other—electronica mixed with acoustic guitar, or the like. Tomas called these playlists "multi-modal"—he reasoned that, in some later layer of the neural network, their incongruities would be sorted out.

Interpretability is a paradoxical concept, because it is inversely related to the work of interpretation. When an output is interpretable, it feels like it requires no interpretation at all—everything fits together, harmonizes, and Tomas and I look at each other and smile, pleased by the apparent coordination of his, my, and the computer's perception. Ironically, it is the apparently "uninterpretable" results that multiply the work of

---

neuronal structure, these parallels have not been borne out by neuroscience research (Edelman 1993, 226-227).

interpretation: when feeding the computer new queries, our attentions range over the results, trying to tune in to what the computer hears. Is it the guitar tone or the rhythm of the drums? A few weeks later, listening to playlists generated by Tomas's algorithm with one of my interlocutors, we try to figure out what the computer is paying attention to in Janet Jackson's 1986 song "Nasty," by listening to the songs it has considered similar. We conclude that it must the distinctive (and now quite dated) sound of the snare drum.

Tomas, the computer, and I make a nice Peircean triad: he directs my attention to features of the music, and I learn to hear them as he has learned to hear them—as we imagine the computer has learned to hear them. I can only adjust my own attention, learning to listen as Tomas has. Tomas, on the other hand, can tweak his system, trying to bring these interpretations into alignment in a feedback loop of which his own, previously trained attention is a crucial part. His sense that the computer should be able to learn how to hear Djent like he has guides his understanding of what the computer can notice, and his own perception is "in the loop."

Listening is thus a common technique for making sense of algorithmic systems: from the ISMIR presenters trying to hear their audio features, to Tomas tuning his neural network, to the ordinary user guessing what about a song led it to be recommended to them. During an interview at a tech company cafeteria, one engineer recounted to me how familiar he had become with the "sound" of his playlisting algorithm: typically, the cafeteria played the algorithmic radio his team had developed, but one day he thought the sequence of songs sounded off ("We would never play that many very popular songs in a row," he told me). He guessed—correctly—that someone had changed to music to a competitor's radio offering.

This interpretation work can be fruitfully thought of in terms of listening—both literally, as in the examples above, and figuratively, in the sense that there is something "listening-like" about these efforts to know. This is an acoustemology of algorithms, a common way of interacting with the outputs of algorithmic systems and attempting to infer their workings. This kind of knowledge about algorithms falls in and out of phase with dominant knowledge practices organized (literally and figuratively) around vision. In other words: "black boxes" sound different than they look.

Exploring the acoustemology of algorithms leads us to a few interesting implications: First, that excess of sensible pattern that musicians call "timbre" and toward which Tomas and I shared our attention plays a significant role. It is hard to describe, but easy to hear—if we were talking about culture, it might be the kind of thing we'd try to capture through "thick description." Second, as listeners oriented to patterns, "insider" engineers and "outsider" users are closer than popularly imagined. In any moderately complex algorithmic system, engineers have no objective or immediate way to see how an output came about. Although they might be able to track down the cause of a given output with enough time, in ordinary day-to-day work, this is prohibitively difficult. This is no immediacy of vision—like outsiders, they range their attention over outputs and attune themselves to patterns. Engineers may differ from users, then, not primarily in what they are able to see, but in how they have learned to hear.

Third, interpretability is a consequence, rather than a cause, of shared attunement. To build an algorithmic system is to stabilize listenings, among members of teams and technical components. And fourth, algorithms are valued over time, through sensory processes that are necessarily interactive, lacking clear distinctions between perceiving subjects and perceived objects. In the course of training his neural net and developing a taste for Djent, Tomas has trained his own attention. He is tuned to the frequencies of chugging guitar riffs and to the meta-patterns of convolutional filters and temporal poolings. These listenings feed back into the working of the algorithm, as it is built. Knock on the black box, hear what it sounds like, tune the parameters, and knock again.

CONCLUSION:
THE ANTHROPOLOGY OF ATTENTION

Open Floor Plan

I am sitting at a desk and my attention is split. The offices of the company I call Whisper have an open floor plan: pods of tables spread across two floors in an old brick building, surrounded by a few meeting rooms. Upstairs is quieter—the office sound system turned down, the average age higher, Sales and HR making phone calls from a few private offices. Downstairs, the engineers. Over the past year, Whisper has grown dramatically, and the pods are packed, Apple laptops plugged into widescreen monitors. Neighbors help out with each other's code, and occasionally, whole pods pick up and move to one of the meeting rooms for their daily "stand-ups" or weekly "retrospectives," where they check in with the team and the product manager responsible for their work. At the end of the floor are a kitchen and a ring of couches, where engineers take breaks from their desks and visitors camp out. Here, the music is louder, and a screen above the couches shows what's playing now and next.

Whisper advertises itself as a "music intelligence" company. "We listen to all the music online," their website claims: this is the life of computer audition outside of the lab, where musical feature extraction is scaled up from small research datasets to the millions of songs available through commercial streaming services and on the web. Whisper uses its own proprietary feature extraction system, which produces a "low-level" representation of sound to be fed into higher-level algorithms that calculate higher-order (more "subjective," I am told) patterns, discerning features like “danceability” and “acousticness.” But this is not the only listening that Whisper performs. Another set of web-crawling bots “listens” to discussions of music that happen online, on blogs, news websites, and the like, analyzing the words that often and distinctively go together to build models of artist similarity. Determining musical similarity is the core of their business, and the work of many Whisper employees is concerned with directing and focusing the attention of their data collection system so that similarities turn out to be meaningful and useful for tasks like making recommendations.

I’ve been visiting for three months as an intern, migrating from desk to desk as employees take their vacations. My current spot is down with the engineers, at the end of the floor near “intern island,” where the summer interns help evaluate the outputs of algorithms that rank popular artists, make playlists, and assess musical similarity. My back is toward most of the 30 or so engineers spread across the room, so although I can’t see them, I can hear them. On my computer, I’m logged in to a company email account, where I chat privately with people sitting around the office. I’m also logged into the company IRC channel—Internet Relay Chat—with most of the employees. The conversation revolves around the currently playing music or the sharing of music-related news links—there’s a new Aphex Twin album, a tech blog just posted a story about a competitor, and so on. Each “squad”—the group of people responsible for a specific software feature like radio or personalization — has its own IRC channel to talk about work and for people from other squads to visit when they have questions. Pods of tables and squads of people are nearly coincident, so these chats tend to include the five or six people sitting near each other, plus a few scattered others. Office conversations are multimodal: a laugh erupts from a pod behind me and a conversation starts out loud regarding something that had just been said in chat, which may have emerged from an email thread including people both in and outside of the local office. As an ethnographer, my view on this situation is partial, as it is for most people in the office — some people do not participate in IRC, email threads only include a few participants, and it is basically impossible to pay attention to every communication channel and still get any work done. As an "intern" outside the usual intern program, I have few responsibilities beside helping with occasional ad hoc projects, and even with all this time to dedicate to observing office communication, the observable seems like the tip of an iceberg. I hear the laughter, I may see the conversation if it's in an IRC channel I am in, but there is always a sense that the conversations I participate in and hear in the office emerge from and regularly dip back into contexts I haven't been privy to.

The open floor plan is a notorious symbol of corporate start-up culture, allowing for agile reconfigurations of office structure, for the display of egalitarianism — very few people have private offices—and for spontaneous, accidental conversations, which are thought to lead to innovation. If the cubicle farm represents the ossified world of old technology companies, ripe for disruption, the open floor plan is the scrappy upstart, with less developed infrastructure, but more powerful for it. But it has its shortcomings. In the interviews I conduct during the workday, in the small conference rooms that surrounded the floor, many people—especially the older employees—remark that the noise of the office can be distracting, and they wear headphones to block out the music and chatter. A scientific literature has emerged that studies the negative effects of ambient noise on cognitive function (e.g. Perham et al. 2013) and open-plan offices on productivity (e.g. Rashid and Zimring 2009, Smith-Jackson and Klein 2008; see Konnikova 2014). Yet many people point out this ambience and the music in particular as essential and valuable elements of the office “culture”—worth preserving even at the cost of productivity—at risk of dissipating as the number of employees grows.

The hustle and bustle of the office, especially in contrast with the relatively sparse flow of fieldwork events while I was not based inside a company, is overwhelming. Interesting conversations and spontaneous meetings have a habit of appearing all of a sudden, as though they precipitate out of the wireless signals carrying conversations around the office. I learn from Ellie, a “data curator” who sits at the desk next to mine, that I can see the digital calendar of anyone in the office for the purposes of scheduling meetings, and this becomes another signal to pay attention to—an imperfect temporal plan of the rooms in which people might be found throughout the day. Whisper’s openness is distributed across the arrangement of its furniture and the arrangement of its digital communications, and it is partial. For the ethnographer, the open floor plan office is thus a blessing and a curse. It offers something of the open village center, where social action is largely public and ethnography makes its historical sense. It overwhelms with potential signals arriving from all directions. But, as in the reality of the idealized fieldwork village, culture, communication, and sociality extend beyond the ethnographer’s field of vision, making a total view impossible as activity moves in and out of sight.

Immersion and the Ethnographic Ear

The open floor plan office, in addition to signaling certain features of the contemporary tech industry and its self-image, embodies the challenges of ethnographic research. Gaining access to the physical space of the office does not guarantee access to all that is going on within it, in spite of its openness.[43] Activity overwhelms the ethnographer's ability to keep track of it all, and it recedes into computer screens and private channels, moving in and out of the field of attention. In *Property, Substance, and Effect*, Marilyn Strathern suggests that this kind of experience is integral to fieldwork's production of knowledge. She recounts a moment of ethnographic attention on Mt. Hagen, Papua New Guinea, catching sight of some men in her peripheral vision:

> I shall never forget my first sight of mounted pearlshells in Mt. Hagen, in 1964, heavy in their resin boards, slung like pigs from a pole being carried between two men, who were hurrying with them because of the weight, a gift of some kind. It was only a glimpse; the men were half-running and their path was almost out of my field of vision. But it belongs to a set of images which have mesmerised me ever since. (Strathern 1999, 8)

This apparently insignificant event would take on tremendous significance for Strathern's research trajectory, as the men carrying a gift crossed over the threshold of her peripheral vision and set off her investigations into the interrelations of gender and gift in Melanesia. "Not to know what one is going to discover," Strathern reflects, "is self-evidently true of discovery" (Strathern 1999, 9), and ethnography's capacity for

[43] Nor does it guarantee that the ethnographer will be able to take anything *out* of it: my access to Whisper was contingent on signing a non-disclosure agreement that allowed me to report non-trade secret details, to be validated by a company representative. Because this approval process turned out to be lengthy and the time available to write this dissertation short, my experiences at Whisper have not been included in this

surprise is facilitated by fieldwork's immersiveness. This immersion, and the experiences of surprise it engenders, is integral to Strathern's understanding of ethnography (see also Guyer 2013 on ethnographic surprise and "the quickening of the unknown"). Ethnographic immersion instrumentalizes the fuzzy boundary of one's ability to see as a mechanism for knowledge production. If other modes of social knowledge production consist in methods that constrain action to a well-defined space within our gaze (the clean borders of surveys, sampling strategies, and structured interviews), ethnography requires immersion such that experience and data overwhelm the senses. The resulting knowledge is necessarily "partial," as Strathern notes in reference to Donna Haraway (1988), and it progresses in unexpected directions as phenomena cross over the threshold of our perception. For Strathern, the sight of the men and the resin boards introduced a new set of research concerns than those she had set out with, and it "dazzled" her, capturing her attention like the Trobriand canoe-prows described by Alfred Gell in "The Technology of Enchantment and the Enchantment of Technology" and the aesthetic traps I discussed in chapter 2 (Gell 1992).

One striking feature of Strathern's account of peripheral vision is how much it seems like *sound*. Immersiveness, surprises that come from the forest, and perceptual thresholds are typical features of discourse about sound, and Strathern reads as though she is grasping at sonic experience through visual discourse. Indeed, we might surmise that what first caught Strathern's attention was not the sight of the pearlshell boards

---

dissertation, although they have helped to train my attention on features of my other fieldwork that might have otherwise escaped notice.

rushing through the woods, but rather their sound, which compelled her to look. We can think of peripheral vision, to bend the senses, as a sonic kind of sight—seeing where your ears are already pointed. Where Strathern characterizes the immersive ethnographic encounter as a technique for "learning to see" (Strathern 2013), we can imagine it instead as a technique for "learning to listen." This process is crucial to ethnography, but also to the developers of recommender systems who, as I described at the end of the previous chapter, learn to listen to the outputs of their systems, bringing their own already-trained ears into feedback relations with algorithms.

"What of the ethnographic ear?", James Clifford asks off-handedly in the introduction to *Writing Culture* (1986: 12), noting the dominance of the "ethnographic gaze" in the anthropological imagination and the relatively untheorized sense of sound. The implication, which has since been picked up in a literature on the anthropology of sound growing in its wake, is that the geometries of sound and vision are divergent: the linear gaze sweeps across a scene like a laser, while the ear is roundly immersed in a sonic surround, to which the listener is vulnerable—ears, as are often noted, cannot be closed. But, as I argued in the last chapter, we should not be too quick to mobilize essentializing ideas about the nature of the senses, and hearing can mean more than bodily presence and vulnerability to vibration. Similarly, phenomena such as peripheral vision draw into question the rigidity of the gaze.

In the offices of Whisper, with my eyes on the computer screen and my ears perked up, ready to receive any noise as signal, I experienced the challenge of developing an ethnographic ear. Sonic environments like these are not unusual for contemporary ethnographic fieldwork, regardless of the site, and they make demands on ethnographers' attention: when we hear so much, what should we listen to? How do we deal with the cocktail party problem I described in the previous chapter—the problem of picking out a particular signal to notice while others clamor for our attention? I propose that we can think of ethnographic immersion sonically as something like being among Helmholtz's "variegated crowd of intersecting wave systems" (1995:57). The practice of ethnographic listening (Forsey 2010), then, is not a matter of discerning a signal buried in noise, but rather of tuning in to signals that exist among a panoply of potential signals. This tuning is socioculturally trained: remember the auditor of Helmholtz's concert scene in the previous chapter distinguishing not only high tones from low, but men from women and music from noise, disaggregating complex signals into socially defined bins.

Although ears are trained, they are constantly retrained in response to the environment in which people hear—we can think of the deep learning programmer learning to hear like his algorithm from the previous chapter, the data gardener learning to distinguish musical "weeds" from desirable growth in the chapter before, or the would-be trapper adjusting his traps in the chapter before that. Or, we might think of the "deep listening" practice of experimentalist and improvisationalist Pauline Oliveros, which "explores the difference between the involuntary nature of hearing and the voluntary, selective nature—exclusive and inclusive—of listening" (Oliveros 1989). Oliveros' close listening and its results recall Strathern's point about scale and ethnographic attention: when we begin to focus on them, "items of knowledge multiply and divide under one's eyes" (Strathern 1999, 8), dissolving under our attentive pressure.

Put this way, Evans-Pritchard's dictum that good fieldworkers must possess a "sensitivity to pattern" (Faubion 2009, 146) takes on a new meaning: not only should ethnographers perform Sahlins' synaptic function, identifying resonant patterns across disparate domains, as described in chapter 3, but they must learn how to listen. The immersion of fieldwork offers an occasion to learn new ways of listening—to educate our attention, becoming sensitized "to certain features of the world," as Tim Ingold puts it (Ingold 2001). In this understanding, the knowledge thus produced "does not lie in the relations between structures in the world and structures in the mind," as an image that can be distorted or inaccurate; instead "the human being [the fieldworker or the engineer] emerges as a centre of awareness and agency whose processes resonate with those of its environment" (Ingold 2001). The capacity to pay attention is a technique of the body (Mauss 1973), thoroughly dependent on the environments through which persons move, but also a potential object of intentional training as people work on the objects of their perception in a process of mutual tuning.

Techniques for Paying Attention

Ethnography, like algorithmic recommendation, is a set of situated techniques for paying attention in immersive conditions. The notion of immersion—the scalar relationship between the ethnographer's limited horizon and the field's overwhelming totality—recalls the notion of information overload, as described in chapter 1. This relationship, understood as a problem, provided the impetus for recommender systems, as techniques for mediating between individuals and archives which they could not possibly draw entirely into their attention. Ethnography and algorithmic recommendation both work in fields that exceed the bounds of perception, facilitating the experience of immersion; that experience of immersion, in turn, makes surprise possible, and it makes surprise—the quickening of the unknown, the crossing of phenomena over the threshold of perception—central to experience. But where the developers of recommender systems have seen a problem, ethnographers see an opportunity. Borrowing Jonathan Sterne's description of the mp3 from chapter 4, one might even say that ethnography is a celebration of the limits of human perception, mobilizing the ethnographer's perceptual horizon as a tool for knowledge production. The distinctiveness of ethnographic knowledge is cut on the edges of the researcher's perception. It is, as Haraway argues and Strathern agrees, through one's partiality that knowledge is produced (Haraway 1988; Strathern 1991).

To understand recommenders as tools for partiality puts them at odds with a popular discourse about big data systems (of which algorithmic recommendation is a typical example) that takes the horizons of human attention to be irrelevant: the overwhelming scale of data supposedly makes it equivalent to the world, and algorithmic processing makes it possible to comprehend all at once (Anderson 2008). However, this understanding of recommenders not as techniques for grasping everything at once but for allocating attention fits better with the actual work of big data in practice, which in spite of this popular discourse is always partial and, like ethnography, exploratory. (And we should not privilege ethnography because it is immersed in "reality" as opposed to big data's immersion in constructed representations—another key to Strathern's theorizing about ethnography is its reliance on inscription and the production of an archive in which the ethnographer can be re-immersed.) These resonances between ethnographic knowledge production and big data have been sounded by a few anthropologists who note similarities in the issues they face regarding scale, dimensionality, and immersion (Gray 2011; Boellstorff 2013; Nelms 2014; Seaver 2015). These similarities, they suggest, point to alternative ways of understanding the relationship between big data and ethnography than the classic confrontation of quantitative and qualitative methods or the "two cultures" of scientific and humanistic inquiry (Snow 1959).

By way of conclusion, I want to return to the arguments made throughout this dissertation, drawing out these themes of attention and technique, which complicate common ideas about the distinctions between human and machine capacities and between taste and technology as ways of ordering experience. I have already suggested that recommender systems are technical configurations for organizing attention, mediating between individuals with limited horizons and archives that exceed them. I have argued that these technical configurations are shaped by ideas about listeners, music, and listening, which consequently come to influence the distribution of attention more generally. This account aligns with broader arguments that a contemporary "attention economy" exists (Simon 1971; Goldhaber 1997; Beller 2007; Crogan and Kinsley 2012) and that the distribution of attention is in critical ways a technical question (Ash 2012; Bucher 2012; Stiegler 2010).[44] The questions that these systems

[44] Attention has also become an explicit object of research interest for machine learning researchers building "attention models," which attempt to break down large analysis tasks and gain efficiency by determining which parts merit the most "attention" from the algorithm: "There are many reasons to be excited about attention. One of them is that attention models simply work better, allowing us to achieve better results with less data. Also, bear in mind that humans clearly have attention. It is something that enables us to

raise are also raised by ethnographic research: How is attention shaped by sociotechnical arrangements? What do we pay attention to? How do we align our attentions with each other's? What techniques do we use to organize our attention?

In the introduction, I argued that it was useful to think of algorithms not as stark, formally defined mathematical operations, but as sociotechnical "algorithmic systems," in which humans play an integral role. This understanding of algorithms runs counter to a popular discourse about music streaming services that pits them against humans; it also counters a tendency in academic criticism that understands the problem with algorithms to be their inhumanness. The humans in the system are responsive to its outputs and capable of reconfiguring it when undesired results come about; their capacity for judgment is a critical element among the many disparate data flows, aggregation processes, and classifying operations that make up an algorithmic system. Their attention to the functioning of the system is learned partially in response to its particularities, partially in their training as engineers, and partially as enculturated persons. The nice thing about working with music, as my interlocutors told me, was that it was easy to perform "sniff tests" at various points to see whether outputs made sense, and these casual evaluations were facilitated by cultural knowledge about music. Thus, the ability to recognize outputs as "correct" was shaped by the musical horizons of people working within the algorithmic system.

---

get results. It's not just an academic concept. If you imagine a really smart system, surely, it, too, will have attention." (http://radar.oreilly.com/2015/08/unsupervised-learning-attention-and-other-mysteries.html)

The heterogeneity of the algorithmic system is the setting in which various techniques bump up against each other: Computational processes depend on data sources that are formatted in a particular way (this is how programs are notoriously “brittle”—they may simply cease to operate if structural requirements are not met), but they also depend on more subtle features of the data flowing through them, like the tendencies of listeners whose behavior may shift over time, or patterns imposed on data by interfaces. In these latter cases, the algorithm does not “break” in the sense that it fails to operate — it continues to produce results, the quality of which becomes a matter of human consideration as parameters are tuned, data sources are reconsidered and rerouted, and the humans in the system attempt to bring the various techniques of the system back into alignment with each other (their own perception included).

This understanding of algorithmic systems does not contradict the common sense of my interlocutors, but is rather in tune with it. Against algorithmic stereotypes, they increasingly emphasize the significance of humans in these systems, rejecting the ideals of objectivity or rationality associated with computing and emphasizing the interplay between them and their human minders, whose cultural competence becomes consequently a matter of more significance. They play in the gaps between supposed human and machine competencies, seeking to disaggregate common sense about them and reimagine how capacities might be distributed across them in the heterogeneous sociotechnical systems in which they work.

With algorithmic systems as the setting, it is easier to see the operations of preferential technics as I described in chapter 1. That these systems blend concerns about circulation and scale with theories about taste and musical preference becomes more obvious when their outputs are understood as a result of human and computer collaboration. Theories about taste and music inform people's work, and this too was not a surprise to my interlocutors, although they tended to understand the variability among people to be a matter of avidity rather than a higher order variability in ideas about the structure of taste.

It is, however, difficult to point to particular theories that inform particular bits of technical infrastructure, aside from bland generalities such as: "people who presume that sound is significant to musical taste will build systems that attempt to take sound into account." While this is to a certain extent true, it obscures the immediately following questions: What is sound? How should it be represented? As described in chapter 4, these questions do not have plainly obvious answers, and sometimes they have surprising ones: At a hackathon in New York City, I met an engineer who was building a system to recommend music on the service Soundcloud, which lets users upload their own music. He wanted a recommender for music that sounded similar, but instead of trying to analyze the sound of the uploaded tracks, he was processing the pattern of users' "likes" on the system, presuming that those would, in aggregate, correspond to their sound. When he generated a playlist for me, it appeared that he was right: we could both hear that the sound of the songs on the playlist was roughly consistent. This was likely some combination of luck, the patterning of taste in individuals, our own efforts to listen for the similarities, and the way that this platform in particular was used at this moment in time, such that "likes" could be used to track apparently specific sonic qualities.

The relationship between theories of taste and technical infrastructures that characterizes preferential technics is not a straightforward mapping, in which a theory is implemented in code and is then sent out into the world to live or die by its correctness or ability to pull human activity in its direction. In practice, it is hard to distinguish theories from techniques—the arrangements of "practical reason," in Mauss's terms—that act and are informed and change in concert with each other. Thus, my attempts to seek out the terms of an emerging preferential technics were not met with a simple statement of theory from my interlocutors in the field. Though I made a habit of asking my interviewees "Why do you think people like the music that they like?", I was routinely answered with "ummm," "uhhh," or a burst of laughter. These theories were not discrete mental representations but ad hoc, contextually specific common senses that depended on the technical situation in which my interlocutors found themselves. It was not simply that technical infrastructures depended on theories of taste, but also the opposite: "theories," insofar as they existed, were manifested in the sets of techniques that came together in the production of algorithmic systems.

Identifying theories of taste that might be straightforwardly analyzed in correspondence to technical infrastructures thus proved impossible. Instead, I examined broadly shared understandings of key elements in music recommendation—listeners, music, and listening—as they manifested in ordinary talk and action, and in technical arrangements. Though, as with any human activity, these understandings might vary locally, the ones outlined here were both common and illustrative of how theory and technique intertwine in algorithmic systems. Thus, in chapter 2, I examined how a common understanding of user variability was mutually constituted with a common way of recording user activity: thinking of music listeners as varying in their avidity for music — their enthusiasm for it—fit with representations of users as aggregations of interaction events, such that more avid listeners would have more interactions. This way of attending to listeners shaped how the developers of recommender systems pursued what had become a standard for evaluating recommender performance: the retention of users. This focus on retention resonated with a classic anthropological example of how culture and technology intermingle—the animal trap. Recommender systems, like traps, are built for capture, and their building is not a strictly technical concern, driven solely by interests like efficacy and efficiency; rather, it is informed by ideas about the entities to be captured, which themselves manifest in the techniques through which the builders of traps build them.

In chapter 3, I turned from listeners to music, outlining an idea common to music recommendation and machine learning more generally: that music (or the objects of machine learning) occupies a space, and that the work of building a recommender system is to make this space navigable. Attending to music as though it constituted a space fit with efforts to arrange it according to similarity, such that similar items and users would be located near each other. However, the abundance of dimensions along which items might be compared required developers to use their enculturated judgment to prune back these dimensions to a smaller set considered salient. These operations shaped the spaces in effectively arbitrary ways, thus contradicting the popular idea that "space" is a neutral grid on which objective differences might be observed. While standard critiques would pit this biasing against a rhetoric of technical objectivity—

these people think they are engaged in an objective mapping exercise, but they are not — many of my interlocutors were well aware of that they influenced these arrangements. So instead, looking for how they understood this influence (and how their understandings might vary from those of outside critics), I found that they used a set of metaphors I call "pastoral," analogizing their role within the algorithmic system to park rangers, gardeners, and farmers. These metaphors locate their work in an apparently "natural," bucolic state, but parks, gardens, and farms are *not* natural. Thus I suggested that these metaphors provide a way between extreme understandings of the music space as either a construction—reflecting the biases and cultural positions of its makers—or an objective discovery—reflecting the "actual" structure of musical similarity. Again, situated understandings emerge that deny a sharp distinction between culture and technology, the human and the algorithm, or bias and objectivity.

The discussion of listening in chapter 4 focused more narrowly on the changing techniques by which computers are being made sensitive to musical sound. Starting from the idea that music (like sound more generally) is essentially mathematical, researchers then face the problem that musical sound cannot be directly fed into an algorithmic system but must be processed first by techniques operating according to some idea of musical salience. Bridging the "semantic gap"—between the numerical representation of sound waves and human understanding of music—is a critical object of research in the academic community. So, although the techniques in question are mathematical operations that reduce the large number of numbers that represent a waveform to a small number that represent its musically salient features, their efficacy is often demonstrated through listening exercises. Counter to an idea that quantification is always a reduction of the sensible world, I suggested that these practices of knowing numbers through listening, or quantitative acoustemology, reflect a two-way resonance between hearing and counting: as sound is attended to mathematically, so numbers are attended to sonically. As researchers develop techniques by which computers pay attention to sound, they try to bring their own attention into alignment with it, and these techniques shape each other.

Anthropology and Cultural Critique

In this dissertation, I have refused several readily available lines of critique against algorithms: I have not argued that technological approaches to understanding culture reduce or misconstrue it; I have not claimed that algorithmic logics per se are essentially ideological, capitalist or neoliberal; and I have not identified algorithms as black boxes that critics must seek to open. Instead, I have suggested that we think of algorithms as heterogeneous and sociotechnical algorithmic systems, which cannot be opened but rather must be traced; which evince multiple coexistent ideas about how the world works; and which do not oppose culture, but are rather constituted by and constitutive of it. This shift in focus does not mean that critique is unimportant or impossible—it directs our attention to aspects of these systems that have been neglected by critiques that presume that we already know what the "problem" is with algorithms.

This approach, as laid out in the end of chapter 1, I take to be quintessentially anthropological. Rather than presuming that I already knew what algorithms were and how they acted, I took that as an empirical question. This approach is advocated by Maurer with regard to number more generally: "We should not fear numbers simply because they are numbers and we think we know what numbers do, always and everywhere," nor should we assume "that whenever we see numbers and math we see something that counts, calculates, equates, desacralizes, and rationalizes" (Maurer 2006, 24). Refusing this common sense about numbers—or algorithms—leaves the outside critic open to surprises in the field and to analyses that open up new possibilities rather than closing them down. Thus, we can find academic research scientists listening to their quantitative models rather than just proving them formulaically, as in chapter 4; we can find developers who think of themselves as farmers and see not just a naturalizing fallacy but a way of figuring complex ecologies of control, as in chapter 3; and we can track the peculiar set of understandings involved in "reducing" users to their interactions with a technical system, as in chapter 2. While the understandings of listeners, music, and listening I have outlined here were prevalent during my fieldwork, they were not the only understandings in play, and in their details there were probably as many theories of taste as there were people.

Recognizing that algorithmic systems are full of people, not only calculations and data sets, has serious consequences for the critique of algorithms. It is key to understanding how these systems actually work, and especially for understanding change over time not as something external to the algorithm, but internal to the algorithmic system. A standard critique might point out that Google's machine learning apparatus has produced an offensive caption for an image, but it cannot tell you anything about how that caption went away after a public outcry. The human minders of algorithmic systems and their culturally trained attentions are critical to these processes and ignored by conventional algorithmic critiques.

The people working in these systems are also much more aware of common critiques than is often assumed. They operate in a cultural milieu where the dominant common sense is that culture and technology are essentially opposed, like humans and machines, and to pursue their work, they have had to reckon with this understanding. They have encountered the tensions of algorithmic systems first-hand, and they are certainly not the "acultural dopes" assumed by critics who think that they can read the contents of a person's mind out of the software they produce. At least in my field sites, they are also familiar with most of the graduate-level critiques levied against their work. However, this does not mean they act on those critiques—or more precisely, it does not mean that they act as critics assume one should act given their critiques.

It should not be surprising that scientists and engineers, like other people, can hold conflicting ideas in their heads at the same time or disregard theoretical arguments in favor of pragmatic action (Deeb and Marcus 2011). However, the unspoken premise of much critique is that, once we highlight these contradictions or dredge them up from the sociotechnical muck, our words will have effects. Either the people working in these systems will have their consciousness raised or, armed with sufficiently theorized critical descriptions, someone else can go and make a change. Though these critiques may note the persistence of human bias in algorithmic systems, they do not take the presence of humans seriously, as people trying to make sense of the world and their own work.

Many of the arguments I have made in this dissertation brought together understandings I encountered in the field with those from my disciplinary training; in sounding out these resonances, I sought to bridge a supposed gap between critical outsiders and uncritical insiders. These resonances provide an opportunity for critics to make their critiques more effective, to treat the people working within algorithmic systems as potential collaborators rather than inaccessible others hidden inside of black boxes and corporations. If we want our critiques to actually make a difference in how these systems are made and in how they work, then we need to engage with the people who make them and make them work.

These people have complex and often ambivalent relationships with the algorithmic systems they operate in. They have diverse understandings of the world that cannot all be read out of interfaces, and they are expert at changing register in discussions of their work—making one set of arguments to investors, another to managers, and another among themselves at the bar after work. If we want to understand how algorithmic systems come to work as they do and to make critiques that affect them, we need to engage with the people within them, in all their complexity and ambivalence.

One night, I was walking to dinner with Richard, one of the first engineers I had met during my fieldwork. “Algorithms should only be used for art,” he told me, seemingly out of nowhere. “I don’t actually think that, but I like how it sounds."

WORKS CITED

Agamben, Giorgio. 2009. *The Signature of All Things: On Method*. Zone Books.

Amatriain, Xavier, Pablo Castells, Arjen de Vries, and Christian Posse. 2012. "Workshop on recommendation utility evaluation: beyond RMSE — RUE 2012." *RecSys '12 Proceedings of the sixth ACM conference on Recommender systems*, 351–352.

Amoore, Louise. 2011. "Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times." *Theory, Culture & Society* 28(6):24–43.

Anderson, C. W. 2012. "Towards a Sociology of Computational and Algorithmic Journalism." *New Media & Society* 15(7):1005–1021.

Anderson, Chris. 2008. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired*. http://www.wired.com/2008/06/pb-theory/

Andrejevic, Mark. 2013. *Infoglut: How Too Much Information is Changing the Way we Think and Know*. New York: Routledge.

Anglade, Amélie. 2014. "Logic-based Modelling of Musical Harmony for Automatic Characterisation and Classification." PhD diss., Queen Mary, University of London.

Antal, Ariane, Michael Hutter, and David Stark, eds. 2015. *Moments of Valuation: Sites of Dissonance*. Oxford: Oxford University Press.

Anzaldúa, Gloria. 1987. *Borderlands/La Frontera: The New Mestiza*. San Francisco: Aunt Lute Books.

Ash, J. 2012. "Attention, Videogames and the Retentional Economies of Affective Amplification." Theory, Culture & Society 29(6):3–26.

Asher, Andrew. 2011 "Search Magic: Discovering How Undergraduates Find Information." Presented at the 2011 American Anthropological Association meeting, Montreal, Canada.

Attali, Jacques. 1985. *Noise: The Political Economy of Music*, translated by Brian Massumi. Minneapolis: University of Minnesota Press.

Balfour, Henry. 1925. "Thorn-Lined Traps and Their Distribution." *Man* 25:33–37.

Balfour, Henry. 1932. "Thorn-Lined Traps in the Pitt-Rivers Museum, Oxford.." *Man* 32:57.

Barad, Karen. 2007. *Meeting the Universe Halfway*. Durham: Duke University Press.

Barlow, John Perry. 1996. "A Declaration of the Independence of Cyberspace." https://projects.eff.org/~barlow/Declaration-Final.html

Beer, David. 2009. "Power through the Algorithm? Participatory Web Cultures and the Technological Unconscious." *New Media & Society* 11(6):985–1002.

Beer, David. 2013. "Algorithms: Shaping Tastes and Manipulating the Circulations of Popular Culture." In *Popular Culture and New Media: The Politics of Circulation*. London: Palgrave McMillan.

Beller, Jonathan. 2007. "Paying Attention." *Cabinet* 24. http://www.cabinetmagazine.org/issues/24/beller.php

Benedict, Ruth. 1934. *Patterns of Culture*. New York: Houghton Mifflin.

Benzecry, Claudio. 2011. *The Opera Fanatic: Ethnography of an Obsession*. Chicago: University of Chicago Press.

Berger, Peter. 1967. *The Social Reality of Religion*. New York: Penguin.

Black, Mary. 1964. "On Formal Ethnographic Procedures." *American Anthropologist* 65:134–135.

Blair, Ann. 2003. "Reading Strategies for Coping with Information Overload Ca. 1550-1700." *Journal of the History of Ideas* 64(1):11–28.

Blair, Ann. 2010. *Too Much to Know: Managing Scholarly Information before the Modern Age*. New Haven: Yale University Press.

Bloor, David. 1976. *Knowledge and social imagery*. Chicago: University of Chicago Press.

Bloor, David. 1982. "Durkheim and Mauss Revisited: Classification and the Sociology of Knowledge." *Studies in the History and Philosophy of Science* 13:267–297.

Boellstorff, Tom. 2005. *The Gay Archipelago: Sexuality and Nation in Indonesia*. Princeton University Press.

Boellstorff, Tom. 2008. *Coming of Age in Second Life*. Princeton: Princeton University Press.

Boellstorff, Tom. 2013. "Making big data, in theory." *First Monday* 18(10).

Borchers, Al, Jon Herlocker, Joseph Konstan, and John Riedl. 1998. "Ganging Up on Information Overload." *Computer* 31(4):106–8.

Born, Georgina. 2005. "On Musical Mediation: Ontology, Technology and Creativity." *Twentieth-Century Music* 2:7–36.

Bourdieu, Pierre. 1972. *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.

Bourdieu, Pierre. 1984. *Distinction*. Cambridge, MA: Harvard University Press.

Bowker, Geoffrey and Susan Leigh Star. 2000. *Sorting Things Out: Classification and its Consequences*. Cambridge: MIT Press.

Bowker, Geoffrey. 2010. "The Ontological Priority of Mediation." In *Melanges for Michel Callon*. Paris: Ecole des Mines.

boyd, danah. 2012. "The Politics of 'Real Names.'" *Communications of the ACM* 55(8):29–31.

Boyer, Dominic. 2008. "Thinking Through the Anthropology of Experts." *Anthropology in Action* 15(2):38–46.

Boyer, Pascal. 1988. *Barricades mystérieuses et Pièges à Penséé: Introduction à l'analyse des épopées Fang*. Paris: Société d'Ethnologie.

Bryson, Bethany. 1996. "'Anything but Heavy Metal': Symbolic Exclusion and Musical Dislikes." *American Sociological Review* 61(5):884–899.

Bucher, Taina. 2012. "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *New Media & Society* 14(7):1164–80.

Bucher, Taina. 2012. "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *New Media & Society* 14(7):1164–80.

Burkart, Patrick and Tom McCourt. 1999. "Infrastructure for the Celestial Jukebox." *Popular Music* 23(3):349–62.

Burkart, Patrick and Tom McCourt. 2006. *Digital Music Wars*. Washington: Rowman & Littlefield.

Burrell, Ian. 2012. "Dr Dre: He's all ears with golden Beats." *Independent*. http://www.independent.co.uk/news/people/profiles/dr-dre-he-s-all-ears-with-golden-beats-8118047.html

Burrell, Jenna. 2009. "The Field Site as a Network: a Strategy for Locating Ethnographic Research." *Field Methods* 21(2):181–99.

Burton, Michael. 1973. "Mathematical anthropology." *Annual Review of Anthropology* 2:189–199.

Bylin, Kyle. 2014. *Promised Land*. BookBaby.

Carr, Nicholas. 2014. *The glass cage: automation and us*. New York: W.W. Norton & Company.

Cat, Jordi. 2014. “Otto Neurath.” *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. http://plato.stanford.edu/archives/win2014/entries/neurath/

Cefkin, Melissa, ed. 2009. *Ethnography and the Corporate Encounter*. Berghahn Books.

Ceglowski, Maciej. 2011. “The Social Graph is Neither.” *Pinboard Blog*. https://blog.pinboard.in/2011/11/the_social_graph_is_neither/

Celma, Òscar. 2010. *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer.

Chan, Anita Say. 2014. *Networking Peripheries*. Cambridge, MA: MIT Press.

Cheney-Lippold, John. 2011. “A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control.” *Theory, Culture & Society* 28(6):164–181.

Chow, Rey. 2012. *Entanglements, or Transmedial Thinking about Capture*. Durham: Duke University Press.

Choy, Timothy. 2011. *Ecologies of Comparison*. Duke University Press.

Clifford, James. 1986. “Introduction.” *Writing Culture*. University of California Press.

Cochoy, Franck. 2007. “A Brief Theory of the ‘Captation’ of Publics: Understanding the Market with Little Red Riding Hood.” *Theory, Culture & Society* 24(7-8):203–23.

Cohen, Lawrence. 1998. *No Aging in India*. University of California Press.

Coote, Jeremy and Anthony Shelton. 1992. *Anthropology, Art, and Aesthetics*. Oxford: Clarendon Press.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest and Clifford Stein. 2009. *Introduction to Algorithms*, third edition. Cambridge, MA: MIT Press and McGraw-Hill.

Crane, Diana. 1972. *Invisible Colleges; Diffusion of knowledge in scientific communities*. Chicago: University of Chicago Press.

Crawford, Kate and Jason Schultz. 2014. “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms.” *Boston College Law Review* 55(1). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2325784

Crawford, Kate. 2014. “I’m noticing a big metaphor gap in how we describe algorithmic filtering. It’s compared to TV networks, newspapers, and used car salesmen.” Tweet. https://twitter.com/katecrawford/status/484755893989683200

Crogan, Patrick, and Samuel Kinsley. 2012. “Paying Attention: Toward a Critique of the Attention Economy.” *Culture Machine* 13.

Crombie, Alistair. 1990. *Science, Optics and Music in Medieval and Early Modern Thought*. Hambledon Press.

Cronon, William. 1996. "Introduction: In Search of Nature." In *Uncommon Ground*, edited by William Cronon. New York: W.W. Norton and Company.

Dahlhaus, Carl. 1983. *Foundations of Music History*, translated by J.B. Robinson. Cambridge: Cambridge University Press.

Darrigol, Olivier. 2007. "The Acoustic Origins of Harmonic Analysis." *Archive for History of Exact Sciences* 61(4):343–424.

Davies, William. 2015. "The Data Sublime." *The New Inquiry*. http://thenewinquiry.com/essays/the-data-sublime/

Deeb, H.N. and George Marcus. 2011. "In the green room: An experiment in ethnographic method at the WTO." *Political and Legal Anthropology Review* 34:51-76.

Deleuze, Gilles. 1992. "Postscript on the Societies of Control." *October* 59: 3–7.

Denning, Peter. 1982. "ACM President's Letter: Electronic Junk." *Communications of the ACM* 25(3):163–65.

Denning, Peter. 2006. "Infoglut." *Communications of the ACM* 49(7):15–19.

DeNora, Tia. 2002. *Music in Everyday Life*. Cambridge, UK: Cambridge University Press.

Derrida, Jacques. 1974. *Of Grammatology*. Johns Hopkins University Press.

Desrosières, Alain. 2012. "Mapping the Social World: From Aggregates to Individuals" *Limn* 2.

Devendorf, Laura and Elizabeth Goodman. 2014. "The Algorithm Multiple, the Algorithm Material." Contours of Algorithmic Life, UC Davis, May.

Diakopoulos, Nicholas. 2013. "Algorithmic Accountability Reporting: On the Investigation of Black Boxes." A Tow/Knight Brief. New York: Columbia Journalism School, Tow Center for Digital Journalism.

Domingos, Pedro. 2015. *The Master Algorithm*. New York: Basic Books.

Dormehl, Luke. 2014. *The Formula: How Algorithms Solve All Our Problems …and Create More*. New York: Penguin.

Dornfeld, Barry. 1998. *Producing Public Television, Producing Public Culture*. Princeton: Princeton University Press.

Douglas, Mary. 1966. *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*. New York: Routledge.

Douglas, Mary. 1972. "Self-Evidence." Proceedings of the Royal Anthropological Institute of Great Britain and Ireland, 1972:27–43.

Dredge, Stuart. 2015. "Spotify Launches Discover Weekly Personalised 'Mixtape' Playlist." *MusicAlly*. http://musically.com/2015/07/20/spotify-discover-weekly-personalised-mixtape-playlist/

Duranti, Alessandro. 2008. "Further Reflections on Reading Other Minds." *Anthropological Quarterly* 81(2):483–494.

Durkheim, Émile and Marcel Mauss. 1963[1903]. *Primitive Classification*. translated by Rodney Needham. Chicago: University of Chicago Press.

Edelman, Gerald. 1993. *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York: Basic Books.

Edmunds, Angela and Anne Morris. 2000. "The Problem of Information Overload in Business Organisations: a Review of the Literature." *International Journal of Information Management* 20(1):17–28.

Elkus, Adam. 2015. "You Can't Handle the (Algorithmic) Truth." *Slate*. http://www.slate.com/articles/technology/future_tense/2015/05/algorithms_aren_t_responsible_for_the_cruelties_of_bureaucracy.html

Ellison, Katherine. 2006. *Fatal News: Reading and Information Overload in Early Eighteenth-Century Literature*. London: Routledge.

Ensmenger, Nathan. 2012. "Is Chess the Drosophila of Artificial Intelligence? A Social History of an Algorithm." *Social Studies of Science* 42(1):5–30.

Eppler, Martin J and Jeanne Mengis. 2004. "The Concept of Information Overload: a Review of Literature From Organization Science, Accounting, Marketing, MIS, and Related Disciplines." *The Information Society* 20(5):325–44.

Erlmann, Veit. 2010. *Reason and Resonance: A History of Modern Aurality*. Zone Books.

Evans-Pritchard, E.E. 1940. *The Nuer: A description of the modes of livelihood and political institutions of a Nilotic people*. Oxford: Clarendon Press.

Evans-Pritchard, E.E. 1976. *Witchcraft, Oracles, and Magic among the Azande*, abridged and with an introduction by Eva Gillies. Oxford: Clarendon Press.

Eyal, Nir. 2013. *Hooked: How to Build Habit-Forming Products*. Self-published, re-published by Portfolio in 2014.

Faubion, James. 2009. “The Ethics of Fieldwork as an Ethics of Connectivity.” In *Fieldwork Is Not What It Used to Be*, edited by James Faubion and George Marcus. Cornell University Press.

Feld, Steven. 1982. *Sound and Sentiment*. University of Pennsylvania Press.

Feld, Steven. 1996. “Waterfalls of Songs: An Acoustemology of Place Resounding in Bosavi, Papua New Guinea.” In *Senses of Place*, edited by Steven Feld and Keith H. Basso. Santa Fe: School of American Research Press: 91-136.

Fenster, Mark. 1991. “The Problem of Taste Within the Problematic of Culture.” *Communication Theory* 1(2):87–105.

Fields, Benjamin. 2011. “Contextualize Your Listening: The Playlist as Recommendation Engine.” PhD diss., Goldsmiths, University of London.

Fischer, Michael. 2009. *Anthropological Futures*. Durham: Duke University Press.

Forsey, Martin Gerard. 2010. “Ethnography as participant listening.” *Ethnography* 11(4):558-572.

Franklin, Seb. 2015. *Control: Digitality as Cultural Logic*. Cambridge, MA: MIT Press.

Fricke, David. 2012. “Jimmy Iovine: The Man With the Magic Ears.” *Rolling Stone*. http://www.rollingstone.com/music/news/jimmy-iovine-the-man-with-the-magic-ears-20120412

Friedersdorf, Conor. 2014. “Would You Rather Get Tips From an Expert of an Algorithm?” *The Atlantic*. http://www.theatlantic.com/business/archive/2014/06/would-you-rather-get-book-advice-from-a-librarian-or-an-algorithm/373650/

Frye, Northrop. 1982. *The Great Code: The Bible and Literature*. New York: Harvest.

Galloway, Alex. 2004. *Protocol: How Control Exists After Decentralization*. Cambridge, MA: MIT Press.

Galloway, Alexander R. 2006. *Gaming: Essays On Algorithmic Culture*. Minneapolis: University Of Minnesota Press.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.

Geertz, Clifford. 1984. “Anti Anti-Relativism.” *American Anthropologist* 86(2):263–278.

Gehl, Robert W. 2014. *Reverse Engineering Social Media: Software, Culture, and Political Economy in New Media Capitalism*. Philadelphia: Temple University Press.

Gell, Alfred. 1992. "The Technology of Enchantment and the Enchantment of Technology." In *Anthropology, Art, and Aesthetics*, edited by Jeremy Coote and Anthony Shelton. Oxford: Oxford University Press, 40–63.

Gell, Alfred. 1996. "Vogel's Net: Traps as Artworks and Artworks as Traps." *Journal of Material Culture* 1(1):15–38.

Gell, Alfred. 1998. *Art and Agency*. Oxford: Oxford University Press.

Gillespie, Tarleton. 2010. "The Politics of 'Platforms.'" *New Media & Society* 12(3):347–64.

Gillespie, Tarleton. 2012. "Can an Algorithm Be Wrong?" *Limn* 1(2).

Gillespie, Tarleton. 2014. "The Relevance of Algorithms." In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo Boczkowski, and Kirsten Foot,. Cambridge, MA: MIT Press, 167-194.

Gitelman, Lisa, ed. 2013. *Raw Data is an Oxymoron*. MIT Press.

Gladwell, Malcolm. 2000. *The Tipping Point*. New York: Little Brown.

Goldberg, David, David Nichols, Brian Oki, and Douglas Terry. 1992. "Using collaborative filtering to weave an information Tapestry." *Communications of the ACM* 35(12):61–71.

Goldhaber, Michael. 1997. "The Attention Economy and the Net." *First Monday* 2(4–7).

Goldstein, Paul. 1994. *Copyright's Highway: The Law and Lore of Copyright from Gutenberg to the Celestial Jukebox*. Stanford: Stanford University Press.

Graebner, Fritz. 1911. *Methode der Ethnologie*. Heidelberg: Carl Winter.

Graham, Paul. 2012. "Startup = Growth." http://www.paulgraham.com/growth.html

Graham, Stephen D. N. 2005. "Software-sorted Geographies." *Progress in Human Geography* 29(5):562–580.

Granka, Laura A. 2010. "The Politics of Search: A Decade Retrospective." *The Information Society* 26(5):364–74.

Gray, Mary. 2011. "Anthropology as BIG DATA: Making the case for ethnography as a critical dimension in media and technology studies." Talk. http://research.microsoft.com/apps/video/default.aspx?id=155639

Grimmelman, James. 2014. "The Facebook Emotional Manipulation Study: Sources." *The Laboratorium*.

http://laboratorium.net/archive/2014/06/30/the_facebook_emotional_manipulation_study_source

Gupta, Akhil and James Ferguson. 1992. "Beyond 'Culture': Space, Identity, and the Politics of Difference." *Cultural Anthropology* 7(1):6–23.

Gurevich, Yuri. 2011. "What is an Algorithm?" Technical Report MSR-TR-2011-116, Microsoft Research. http://research.microsoft.com/pubs/155608/209-3.pdf

Gusterson, Hugh. 1996. *Nuclear Rites*. Berkeley: University of California Press.

Gusterson, Hugh. 1997. "Studying Up Revisited." PoLAR: Political and Legal Anthropology 20(1):114–119.

Guyer, J I, N Khan, J Obarrio, C Bledsoe, J Chu, S Bachir Diagne, K Hart, et al. 2010. "Introduction: Number as Inventive Frontier." *Anthropological Theory* 10(1-2):36–61.

Guyer, Jane. 2013. "'The Quickening of the Unknown': Epistemologies of Surprise in Anthropology." *HAU: Journal of Ethnographic Theory* 3(3):283–307.

Hacking, Ian. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.

Hacking, Ian. 1991. "The Making and Molding of Child Abuse." *Critical Inquiry*. 17(2):253–288.

Haimson, Oliver, Jed Brubaker, Lynn Dombrowski, and Gillian Hayes. 2015. "Disclosure, Stress, and Support During Gender Transition on Facebook." In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1176–1190.

Hallinan, Blake and Ted Striphas. 2014. "Recommended for You: the Netflix Prize and the Production of Algorithmic Culture." *New Media & Society*. Available online before print.

Halloy, Arnaud. 2015. "Pierre Smith's Mind Traps: the Origin and Ramifications of a Theory of Ritual Efficacy." *Anthropological Theory* 15(3):358–74.

Hannerz, Ulf. 2003. "Being There… and There… and There!" *Ethnography* 4(2):201–16.

Haraway, Donna Jeanne. 1994. "A Game of Cat's Cradle: Science Studies, Feminist Theory, Cultural Studies." *Configurations* 2(1):59–71.

Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14(3):575:599.

Harding, Sandra. 1986. *The Science Question in Feminism*. Ithaca: Cornell University Press.

Harding, Sandra. 1995. "'Strong objectivity': A response to the new objectivity question." *Synthese* 104(3):331–349.

Harvey, Eric. 2013. "I Started a Joke: 'PBR&B' and What Genres Mean Now." *Pitchfork*, October 7. http://pitchfork.com/thepitch/95-i-started-a-joke-the-saga-of-pbrb/

Harvey, Eric. 2014. "Station to Station: The Past, Present, and Future of Streaming Music." *Pitchfork*. http://pitchfork.com/features/cover-story/reader/streaming/

Helmholtz, Hermann von. 1995. "On the Physiological Causes of Harmony in Music." In *Science and Culture: Popular and Philosophical Essays*, edited by David Cahan. Chicago: University of Chicago Press. 46-75.

Helmreich, Stefan. 1998. "Recombination, Rationality, Reductionism and Romantic Reactions: Culture, Computers, and the Genetic Algorithm." *Social Studies of Science* 28(1):39–71.

Helmreich, Stefan. 2003. "Trees and seas of information: Alien kinship and the biopolitics of gene transfer in marine biology and biotechnology." *American Ethnologist* 30(3):340–358.

Helmreich, Stefan. 2007. "An anthropologist underwater: Immersive soundscapes, submarine cyborgs, and transductive ethnography." *American Ethnologist* 34(4):621–641.

Helmreich, Stefan. 2009. *Alien Ocean*. University of California Press.

Helmreich, Stefan. 2010. "Listening Against Soundscapes." *Anthropology News* December 2010:10.

Helmreich, Stefan.1998. *Silicon Second Nature: Culturing Artificial Life in a Digital World*. Berkeley: University of California Press.

Hennion, Antoine. 2001. "Music Lovers: Taste as Performance." *Theory, Culture, and Society* 18(5):1–22.

Hennion, Antoine. 2004. "Pragmatics of Taste." *The Blackwell Companion to the Sociology of Culture*. Wiley Online Library, 131–44.

Hennion, Antoine. 2007. "Those things that hold us together: Taste and sociology." *Cultural Sociology* 1(1):97–114.

Herlocker, Jonathan L, Joseph A Konstan, Loren G Terveen, and John T Riedl. 2004. "Evaluating Collaborative Filtering Recommender Systems." *ACM Transactions on Information Systems* 22(1):5–53.

Highmore, Ben. 2010. "Bitter after Taste: Affect, Food, and Social Aesthetics. In *The Affect Theory Reader*, edited by Melissa Greg and Gregory Seigworth. Durham: Duke University Press.

Hill, Kashmir. 2011. “Resisting The Algorithms.” *Forbes Tech*. http://www.forbes.com/sites/kashmirhill/2011/05/05/resisting-the-algorithms/

Hockey, Jenny. 2002. “Interviews as Ethnography? Disembodied Social Interaction in Britain.” In *British Subjects: An Anthropology of Britain*, edited by Nigel Rapport. Oxford: Berg, 209–222.

Hodder, Ian. 2012. Entangled:An Archaeology of the Relationships between Humans and Things. Sussex: Wiley-Blackwell.

Holmes, Douglas and George Marcus. 2005. “Cultures of expertise and the management of globalisation: towards a re-functioning of ethnography.” In *Global assemblages: technology, politics and ethics as anthropological problems*, edited by Aihwa Ong & Stephen Collier. Oxford/New York: Blackwell, 235–252.

Houghton, Bruce. 2013. “Beats Music: Daisy Gets A Name, Hiring More Staff.” *Hypebot*. http://www.hypebot.com/hypebot/2013/08/beats-music-daisy-gets-a-name-hiring-more-staff.html

Hrdlička, Aleš. 1916. “Indian Trap Pits along the Missouri.” *American Anthropologist* 18(4):546–547.

Ingold, Tim. 1974. “On Reindeer and Men.” *Man* 9(4):523–538.

Ingold, Tim. 1988. *Hunters, Pastoralists and Ranchers: Reindeer Economies and their Transformations*. Cambridge: Cambridge University Press.

Ingold, Tim. 2001. “Beyond Art and Technology: The Anthropology of Skill.” In *Anthropological Perspectives on Technology,* edited by Michael Schiffer. Albuquerque: University of New Mexico Press, 17–32.

Ingold, Tim. 2001. “From the transmission of representations to the education of attention.” In *The debated mind: Evolutionary psychology versus ethnography,* edited by Harvey Whitehouse. New York: Berg, 113–153.

Ingold, Tim. 2008. “When ANT meets SPIDER: Social theory for arthropods.” In *Material Agency*, edited by Carl Knappett, and Lambros Malafouris. New York: Springer.

Introna, Lucas D., and Helen Nissenbaum. 2000. “Shaping the Web: Why the Politics of Search Engines Matters.” *The Information Society* 16(3):169–185.

Iyengar, Sheena and Mark Lepper. 2000. “When Choice is Demotivating: Can One Desire Too Much of a Good Thing?” *Journal of Personality and Social Psychology* 79(6):995–1006.

Jackson, Myles. 2006. *Harmonious Triads*. MIT Press.

James, Robin. 2014. “An attempt at a precise & substantive definition of ‘neoliberalism,’ plus some thoughts on algorithms.” *Cyborgology*. http://thesocietypages.org/cyborgology/2014/07/19/an-attempt-at-a-precise-substantive-definition-of-neoliberalism-plus-some-thoughts-on-algorithms/

Jameson, Fredric. 1991. *Postmodernism, or The Cultural Logic of Late Capitalism*. Durham, NC: Duke University Press, 1991.

Jayne, Caroline Furness. 1906. *String Figures and How to Make Them*. New York: Charles Scribner’s Sons.

Jenkins, Henry, Sam Ford, and Joshua Green. 2013. *Spreadable Media: Creating Value and Meaning in a Networked Culture*. New York: NYU Press.

Jennings, David. 2007. *Net, Blogs and Rock ’n’ Roll*. Boston: Nicholas Brealey.

Jensen, Casper B. 2010. “Asymmetries of Knowledge: Mediated Ethnography and ICT for Development.” *Methodological Innovations Online* 5(1):72–85.

Jones, Graham M. 2014. “Secrecy.” *Annual Review of Anthropology* 43(1):53–69.

Kant, Immanuel. 1914[1790]. *Critique of Judgment*, translated by J.H. Bernard. London: Macmillan.

Kavanagh, Donncha. 2010. “Reviewing March’s Vision.” Paper submitted to the 28th Standing Conference on Organisational Symbolism, Lille.

Kimmel, Eric and Trina Hyman. 1994. *Hershel and the Hanukkah Goblins*. Holiday House.

Kittler, Friedrich. 2006. “From Discourse Networks to Cultural Mathematics: An Interview with Friedrich A. Kittler.” Interviewed by John Armitage. *Theory, Culture, and Society* 23(7-8):17–38.

Knijnenburg, Bart, Martijn Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. “Explaining the user experience of recommender systems.” *User Modeling and User-Adaptive Interfaces* 22:441–504.

Kockelman, Paul. 2013. “The Anthropology of an Equation: Sieves, Spam Filters, Agentive Algorithms, and Ontologies of Transformation.” *HAU: Journal of Ethnographic Theory* 3(3):33–61.

Konnikova, Maria. 2014. “The Open-Office Trap.” *The New Yorker*. http://www.newyorker.com/business/currency/the-open-office-trap

Konstan, Joseph and John Riedl. 2012. “Deconstructing Recommender Systems.” *IEEE Spectrum*. http://spectrum.ieee.org/computing/software/deconstructing-recommender-systems

Koren, Yehuda. 2008. “Factorization Meets the Neighborhood: a Multifaceted Collaborative Filtering Model.” *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 426–34.

Kowalski, Robert. 1979. “Algorithm = Logic + Control.” *Communications of the ACM* 22(7):424–436.

Kroeber, Alfred and Clyde Kluckhohn. 1952. *Culture: A Critical Review of Concepts and Definitions*. New York: Vintage Books.

Krögel, Alison. 2010. “Quechua Sheepherders on the Mountain Plains of Wyoming: The (In)hospitality of U.S. Guest Worker Programs.” *The Journal of Latin American and Caribbean Anthropology* 15(2):261–288.

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lansing, Stephen. 1991. *Priests and Programmers: Technologies of Power in the Engineered Landscape of Bali*. Princeton: Princeton University Press.

Latour, Bruno and Steve Woolgar. 1979. *Laboratory Life: The Social Construction of Scientific Facts*. London: Sage.

Latour, Bruno. 1987. *Science in Action*. Cambridge, MA: Harvard University Press.

Latour, Bruno. 2010. “Tarde’s Idea of Quantification.” In *The Social After Gabriel Tarde: Debates and Assessments*, ed. Matei Candea. New York: Routledge, 145-162.

Lemonnier, Pierre. 1992. *Elements for an Anthropology of Technology*. University of Michigan Press.

Lemonnier, Pierre. 2012. *Mundane Objects: Materiality and Non-verbal Communication*. San Francisco: Left Coast Press.

Lena, Jennifer. 2012. *Banding Together: How Communities Create Genres in Popular Music*. Princeton: Princeton University Press.

Lévi-Strauss, Claude. 1963. “The Bear and the Barber.” *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*. 93(1):1–11.

Lévi-Strauss, Claude. 1966. *The Savage Mind*. Chicago: University of Chicago Press.

Lévi-Strauss, Claude. 1971. *Totemism,* translated by Rodney Needham. Boston: Beacon Press.

Lien, Marianne. 1997. *Marketing and Modernity*. London: Berg.

Lizardo, Omar. 2014. "Taste and the Logic of Practice in Distinction." *Czech Sociological Review* 50(3):335–364.

Loeb, Shoshana and Douglas Terry. 1992. "Information Filtering." *Communications of the ACM* 35(12):26–28.

Logan, Beth. 2000. "Mel Frequency Cepstral Coefficients for Music Modeling." *International Symposium on Music Information Retrieval* 28:5.

Lury, Celia. 2012. "Going Live: Towards an Amphibious Sociology." *The Sociological Review* 60:184–97.

Lyon, Stephen. 2003. *Surveillance as Social Sorting*. New York: Routledge.

MacKenzie, Donald. 2006. *An Engine, Not a Camera: how financial models shape markets*. Cambridge, MA: The MIT Press.

Mackenzie, Donald. 2014. "A Sociology of Algorithms: High-Frequency Trading and the Shaping of Markets." Working paper.

Mager, Astrid. 2012. "Algorithmic Ideology: How Capitalist Society Shapes Search Engines." *Information, Communication & Society* 15(5):769–787.

Mahmud, Lilith. 2012. "'The World Is a Forest of Symbols': Italian Freemasonry and the Practice of Discretion." *American Ethnologist* 39(2):425–38.

Mahnke, Martina and Emma Uprichard. 2014 "Algorithming the Algorithm." In *Society of the Query Reader: Reflections on Web Search,* edited by René König and Miriam Rasch. Amsterdam: Institute of Network Cultures.

Malinowski, Bronisław. 1922. *Argonauts of the Western Pacific*. London: Routledge & Kegan Paul, ltd.

Marcus, George and Fred Myers. 1995. *The Traffic in Culture: Refiguring Art and Anthropology*. Berkeley: University of California Press.

Marcus, George and Michael Fischer. 1986. *Anthropology as Cultural Critique*. Chicago: University of Chicago Press.

Marcus, George E. 1995. "Ethnography in/of the World System: the Emergence of Multi-Sited Ethnography." *Annual Review of Anthropology* 24:95–117.

Marcus, George. 1997. "The Uses of Complicity in the Changing Mise-en-Scène of Anthropological Fieldwork." *Representations* 59:85-108.

Margolis, Jane and Allan Fisher. 2002. *Unlocking the Clubhouse: Women in Computing*. Cambridge, MA: MIT Press.

Markham, Annette. 1998. *Life Online: Researching real experiences in virtual space*. Walnut Creek, CA: AltaMira Press.

Martin, Andrew. 2011. "Headphones With Swagger (and Lots of Bass)." *The New York Times*. http://www.nytimes.com/2011/11/20/business/beats-headphones-expand-dr-dres-business-world.html

Martin, Emily. 1998. "Anthropology and the Cultural Study of Science." *Science, Technology & Human Values* 23(1):24–44.

Mason, Otis T. 1900. "Traps of the Amerinds—a Study in Psychology and Invention." *American Anthropologist* 2(4):657–75.

Massanari, Adrienne. 2010. "Designing for Imaginary Friends: Information Architecture, Personas and the Politics of User-Centered Design." *New Media & Society* 12(3):401–16.

Maurer, Bill. 2005. *Mutual Life, Limited: Islamic Banking, Alternative Currencies, Lateral Reason*. Princeton: Princeton University Press.

Maurer, Bill. 2006. "The Anthropology of Money." *Annual Review of Anthropology* 35(1):15–36.

Mauss, Marcel. 1973[1934]. "Techniques of the body." *Economy and Society* 2(1):70–88.

Mauss, Marcel. 1990[1950]. *The Gift: The Form and Reason for Exchange in Archaic Societies,* translated by W.D. Halls. New York: W.W. Norton.

McDonnell, Terence E. 2014. "Drawing Out Culture: Productive Methods to Measure Cognition and Resonance." *Theory and Society* 43 (3-4):247–74.

McFee, Brian. 2012. "More like this: machine learning approaches to music similarity." PhD diss., University of California, San Diego.

McKelvey, Fenwick. 2014. "Algorithmic Media Need Democratic Methods: Why Publics Matter." *Canadian Journal of Communication* 39(4).

Mead, Margaret. 1935. *Sex and Temperament*. New York: Harper Perennial.

Metzger, Duane and Gerald Williams. 1963. "A Formal Ethnographic Analysis of Tenejapa Ladino Weddings." *American Anthropologist*. 65(5):1076–1101.

Milgram, Stanley. 1970. "The Experience of Living in Cities." *Science* 167:1461–1468.

Miller, Daniel. 2000. "The Fame of Trinis: Websites as Traps." *Journal of Material Culture* 5(1):5–24.

Mills, Mara. 2010. “Deaf Jam: From Inscription to Reproduction to Information.” *Social Text* 28(1):35-58.

Moggridge, Bill. 2007. *Designing Interactions*. Cambridge, MA: MIT Press.

Mol, Annemarie. 2002. *The Body Multiple: Ontology in Medical Practice*. Durham: Duke University Press.

Mookerji, R.R. 1939. “A Nicobarese Rat-Trap.” *Man* 39:8–9.

Morris, Jeremy. 2015. “Curation by code: Infomediaries and the data mining of taste.” *European Journal of Cultural Studies* 18(4-5):446–463.

Mosco, Vincent. 2004 *The Digital Sublime: Myth, Power, and Cyberspace*. Cambridge, MA: MIT Press.

Nakamura, Lisa. 2009. “The Socioalgorithmics of Race: Sorting it Out in Jihad Worlds.” In *New Media and Surveillance,* edited by Kelly Gates and Shoshana Magnet. New York: Routledge.

Napoli, Philip M. 2014. “On Automation in Media Industries: Integrating Algorithmic Media Production into Media Industries Scholarship.” *Media Industries* 1(1).

Nelms, Taylor. 2014. “Immersion: Four Beginnings for an Anthropology of Big Data.” https://medium.com/@tnelms/immersion-4c103c0fcac5

Nelson, Diane. 2009. *Reckoning*. Duke University Press.

Nielsen. 2014. “2014 Nielsen Music U.S. Report.”

Nilsson, Johan. 2013. “Studying consumers? Recent social science interest in market research.” Working paper, Linköping University.

Norton, Quinn. 2013. “ACM Web Science talk, as written.” *Quinn Said.* http://www.quinnnorton.com/said/?p=721

Oliveros, Pauline. 1989. *Deep Listening with the Deep Listening Band*. New Albion. Cassette tape. http://deeplistening.org/site/

Orr, Julian. 1996. *Talking about Machines: An Ethnography of a Modern Job*. Ithaca: Cornell University Press.

Ortner, Sherry. 1995. “Resistance and the Problem of Ethnographic Refusal.” *Comparative Studies in Society and History* 37(1):173–93.

Oudshoorn, Nelly, Els Rommes, and Marcelle Stienstra. 2004. “Configuring the User as Everybody: Gender and Design Cultures in Information and Communication Technologies.” *Science, Technology & Human Values* 29(1):30–63.

Paasonen, Susans. 2011. *Carnal Resonance*. MIT Press.

Paerregaard, Karsten. 2003. "The Resonance of Fieldwork. Ethnographers, Informants and the Creation of Anthropological Knowledge." *Social Anthropology* 10(3):319–334.

Palermino, Chris Leo. 2014. "Forgotify Plays Spotify's 4 Million Unheard Songs." *Billboard*. http://www.billboard.com/biz/articles/news/digital-and-mobile/5901260/forgotify-plays-spotifys-4-million-unheard-songs

Pariser, Eli. 2011. *The Filter Bubble: What the Internet is Hiding from You*. New York: Penguin.

Parisi, Luciana. 2013. *Contagious Architecture*. Cambridge, MA: MIT Press.

Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press.

Peate, Iorwerth C. 1934. "Severn Eel-Traps." *Man* 34:153–54.

Perham, N., Hodgetts, H., and Banbury, S. 2013. "Mental arithmetic and non-speech office noise: an exploration of interference-by-content." *Noise Health*15(62):73-8.

Peters, John Durham. 2004. "Helmholtz, Edison, and Sound History." In *Memory Bytes: History, Technology, and Digital Culture*, edited by Rabinovitz and Geil. Durham: Duke University Press, 177-198.

Peters, John Durham. 2015. *The Marvelous Clouds*. Chicago: University of Chicago Press.

Peterson, Richard and Robert Kern. 1996. "Changing highbrow taste: From snob to omnivore." *American Sociological Review* 61(5):900–907.

Pfaffenberger, Bryan. 1992. "Social Anthropology of Technology." *Annual Review of Anthropology* 21:491–516.

Pham, Alex. 2012. "Discovery: TheKey to Digital Fortune." *FutureSound White Paper: Why Discovery, Data and Licensing Are Key to the New Music Business*. Billboard.

Phillips, P. Jonathon, Fang Jiang, Abhijit Narvekar, Julianne Ayyad, and Alice O'Toole. 2011. "An other-race effect for face recognition algorithms." *ACM Transactions on Applied Perception* 8(2):14.

Pierce, Jennifer. 1995. "Reflections on Fieldwork in a Complex Organization: Lawyers, Ethnographic Authority, and Lethal Weapons. In *Studying Elites Using Qualitative Methods*, edited by Rosanna Herz and Jonathan Imber. London: Sage.

Pinch, Trevor and Wiebe Bijker. 1984. "The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other." *S19ocial Studies of Science* 14(3):399–441.

Pinker, Steven. 1997. *How the Mind Works*. New York: W.W. Norton.

Popper, Ben. 2014. "How Beats conquered the world." *The Verge*. http://www.theverge.com/2014/6/12/5800966/beats-powerbeats-iovine-dre-apple-headphones

Powers, Devon. 2014. "Lost in the Shuffle: Technology, History, and the Idea of Musical Randomness." *Critical Studies in Media Communication* 31(3):244–264.

Puschmann, Cornelius and Jean Burgess. 2014. "Metaphors of Big Data." *International Journal of Communication* 8:1690–1709.

Rappaport, Roy. 1968. *Pigs for the Ancestors: Ritual in the Ecology of a New Guinea People*. New Haven: Yale University Press.

Rashid, Mahbub and Craig Zimring. 2008. "A Review of the Empirical Literature on the Relationships Between Indoor Environment and Stress in Health Care and Office Settings: Problems and Prospects of Sharing Evidence." *Environment and Behavior March* 40(2):151-190.

Reinert, Hugo and Tor A. Benjaminsen. 2015. "Conceptualising Resilience in Norwegian Sámi Reindeer Pastoralism." *Resilience* 3(2):95–112.

Reinert, Hugo. 2008. "The corral and the slaughterhouse." PhD diss., University of Cambridge.

Resnick, Paul, Neophytos Iacovou, Miitesh Suchak, Peter Bergstrom, and John Riedl. 1994. "GroupLens: an Open Architecture for Collaborative Filtering of Netnews." *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work,* 175–86.

Riedl, John and Joseph Konstan. 2002. *Word of Mouse: The Marketing Power of Collaborative Filtering*. New York: Warner Books.

Riles, Annelise. 2001 "Encountering Amateurism: John Henry Wigmore and the Uses of American Formalism." *Public Law and Legal Theory Papers*. Working Paper 39.

Riles, Annelise. 2004. "Real time: Unwinding technocratic and anthropological knowledge." *American Ethnologist* 31:392-405.

Riles, Annelise. 2011. *Collateral Knowledge: Legal Reasoning in the Global Financial Markets*. Chicago: University of Chicago Press.

Romney, A. Kimball, Roger Shepard, and Sara Nerlove, eds. 1972. *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences: Applications*, Volume 2. New York: Seminar Press.

Sahlins, Marshall. 1976. *Culture and Practical Reason*. Chicago: University of Chicago Press.

Salmon, Felix. 2014. "Netflix's dumbed-down algorithms." *Reuters blog*, January 3. http://blogs.reuters.com/felix-salmon/2014/01/03/netflixs-dumbed-down-algorithms/

Samuels, David W, Louise Meintjes, Ana Maria Ochoa, and Thomas Porcello. 2010. "Soundscapes: Toward a Sounded Anthropology." Annual Review of Anthropology 39(1):329–45.

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms." *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*. 64th Annual Meeting of the International Communication Association. May 22, 2014, Seattle, WA, USA.

Sandvig, Christian. 2014. "Corrupt Personalization." *Social Media Collective Blog*. http://socialmediacollective.org/2014/06/26/corrupt-personalization/

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt Brace.

Schlanger, Nathan. 2006. *Marcel Mauss: Techniques, Technology and Civilization*. New York: Berghahn Books.

Schüll, Natasha Dow. 2012. *Addiction by Design: Machine Gambling in Las Vegas*. Princeton: Princeton University Press.

Schwartz, Barry. 2004. *The Paradox of Choice: Why More is Less*. New York: Ecco.

Seaver, Nick. 2011. "'This Is Not a Copy': Mechanical Fidelity and the Re-enacting Piano." *differences* 22(2-3):54–73.

Seaver, Nick. 2012. "Algorithmic Recommendations and Synaptic Functions." *Limn* 2.

Seaver, Nick. 2014. Bastard Algebra. In *Data! Now Bigger and Better*. Prickly Paradigm Press.

Seaver, Nick. 2015. "The Nice Thing About Context Is That Everyone Has It." *Media, Culture & Society* 37(7):1101–9.

Shapin, Steven. 2012. "The Sciences of Subjectivity." *Social Studies of Science* 42(2):170–184.

Shapin, Steven. 2012. “The Tastes of Wine: Towards a Cultural History.” *Rivista de Estetica*.

Shardanand, Upendra and Pattie Maes. 1995. “Social Information Filtering: Algorithms for Automating ‘Word of Mouth.’” *CHI ’95*.

Shardanand, Upendra. 1994. “Social information filtering for music recommendation.” Masters thesis, Massachusetts Institute of Technology.

Simon, Herbert A. 1971. "Designing Organizations for an Information-Rich World.” In *Computers, Communication, and the Public Interest*, edited by Martin Greenberger. Baltimore, MD: The Johns Hopkins Press.

Singer, Natasha. 2014. “Whether Working or Job Seeking, the Algorithm Is Watching.” *New York Times*, December 28. http://bits.blogs.nytimes.com/2014/12/28/whether-working-or-job-seeking-the-algorithm-is-watching/

Smith-Jackson, Tonya and Katherine Klein. 2009. “Open-plan offices: Task performance and mental workload.” *Journal of Environmental Psychology* 29(2): 279–289.

Smith, Pierre. 1984. “Le ‘Mystère’ et ses masques chez les Bedik.” *L’Homme* 24(3-4):5–33.

Snider, Laureen. 2014. “Interrogating the Algorithm: Debt, Derivatives and the Social Reconstruction of Stock Market Trading.” *Critical Sociology* 40(5):747–61.

Snow, C.P. 1959. *The Two Cultures and the Scientific Revolution*. New York: Cambridge University Press.

Stallabrass, Julian. 2007. “What's in a Face? Blankness and Significance in Contemporary Art Photography.” *October* 122(9):71–90.

Stefflre, Volney. 1965. “Simulation of People’s Behavior Toward New Objects and Events.” *American Behavioral Scientist* 8(9):12–15.

Stefflre, Volney. 1968. *New Products and New Enterprises: A Report on an Experiment in Applied Social Science*. Market Structure Studies, Inc.

Steiner, Christopher. 2012. *Automate this: How algorithms came to rule our world*. New York: Portfolio.

Sterne, Jonathan. 2003. “Bourdieu, Technique and Technology.” *Cultural Studies* 17(3-4):367–89.

Sterne, Jonathan. 2003. *The Audible Past*. Durham, NC: Duke University Press.

Sterne, Jonathan. 2006. "The Mp3 as Cultural Artifact." *New Media & Society* 8(5):825–42.

Sterne, Jonathan. 2012. *MP3: The Meaning of a Format*. Durham, NC: Duke University Press.

Stiegler, Bernard. 1998. *Technics and Time 1, The Fault of Epimethus*. Stanford: Stanford University Press.

Stiegler, Bernard. 2010. *Taking Care of Youth and the Generations*, translated by S. Barker, Stanford, CA: Stanford University Press.

Stocking, George. 1982. "Franz Boas and the Culture Concept in Historical Perspective." In *Race, Culture, and Evolution: Essays in the History of Anthropology*. Chicago: University of Chicago Press, 195–233.

Stoller, Paul. 1989. *The Taste of Ethnographic Things: The Senses in Anthropology*. Philadelphia: University of Pennsylvania Press.

Stowell, Dan and Mark Plumbley. "Acoustic detection of multiple birds in environmental audio by Matching Pursuit." Technical note, ICML 2013 Bird Challenge.

Strathern, Marilyn. 1991. *Partial Connections*. Walnut Creek: AltaMira Press.

Strathern, Marilyn. 1995. "The nice thing about culture is that everyone has it." In *Shifting Contexts: Transformations in Anthropological Knowledge*. London: Routledge, 153–176.

Strathern, Marilyn. 1999. *Property, substance, and effect*. The Athlone Press.

Strathern, Marilyn. 2013. *Learning to see in Melanesia*. Masterclass Series 2. Manchester: HAU Society for Ethnographic Theory.

Striphas, Ted. 2015. "Algorithmic Culture." *European Journal of Cultural Studies* 18(4-5):395–412.

Suchman, Lucy. 1996. "Constituting shared workspaces." In *Cognition and Communication at Work*, edited by D. Middleton and Y. Engestrom. Cambridge: Cambridge University Press.

Sunstein, Cass. 2015. *Choosing Not to Choose: Understanding the Value of Choice*. London: Oxford University Press.

Swancutt, Katherine. 2012. "The captive guest: spider webs of hospitality among the Nuosu of Southwest China." *Journal of the Royal Anthropological Institute* 18(S1):S103–S116.

Szwed, John. 2010. *Alan Lomax: The Man Who Recorded the World*. New York: Viking.

Tarde, Gabriel. 1903. *The Laws of Imitation*, translated by E.C. Parsons. New York: Henry Holt and Company.

Tarde, Gabriel. 1969. "Basic Principles." In *On Communication and Social Influence*, edited by Terry N. Clark. Chicago: University of Chicago Press, 143-148.

Thaler, Richard and Cass Sunstein. 2006. *Nudge*. New York: Penguin.

Thrift, Nigel. 2005. *Knowing Capitalism*. London: Sage.

Titlow, John Paul. 2013. "Screw Algorithms! The New Music Service From Beats Uses Celebrity Curators To Show You New Music." *Fast Company*. http://www.fastcolabs.com/3019830/screw-algorithms-the-new-music-service-from-beats-uses-celebrity-curators-to-show-you-new-mu

Toffler, Alvin. 1970. *Future Shock*. Bantam.

Traweek, Sharon. 1988. *Beamtimes and Lifetimes: The World of High Energy Physics*. Cambridge: Harvard University Press.

Trower, Shelley. 2011. *Senses of Vibration: A History of the Pleasure and Pain of Sound*. Bloomsbury.

Tsing, Anna. 2004. *Friction*. Princeton University Press.

Tufekci, Zeynep. 2014. "The Year We Get Creeped Out by Algorithms." *NiemanLab Predictions for Journalism 2015*. http://www.niemanlab.org/2014/12/the-year-we-get-creeped-out-by-algorithms/

Tufekci, Zeynep. 2014. "What Happens to #Ferguson Affects Ferguson: Net Neutrality, Algorithmic Fitlering and Ferguson." *Medium*. https://medium.com/message/ferguson-is-also-a-net-neutrality-issue-6d2f3db51eb0

Turow, Joseph. 2011. *The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth*. New Haven: Yale University Press.

TV Tropes. "Clutching Hand Trap." http://tvtropes.org/pmwiki/pmwiki.php/Main/ClutchingHandTrap

Tyler, Stephen. 1986. "Post-Modern Ethnography: From Document of the Occult to Occult Document." In *Writing Culture*. University of California Press.

Umesh, S., L. Cohen, and D. Nelson. "Fitting the Mel scale." *Acoustics, Speech, and Signal Processing*, 1999.

Urbano, Julián, Dmitry Bogdanov, Perfecto Herrera, Emilia Gómez and Xavier Serra. 2014. What is the Effect of Audio Quality on the Robustness of MFCCs and

Chroma Features? *International Symposium on Music Information Retrieval* 2014.

Uricchio, William. 2011. "The Algorithmic Turn: Photosynth, Augmented Reality and the Changing Implications of the Image." *Visual Studies* 26(1):25–35.

Van Buskirk, Eliot. 2013. "Project 'Daisy' Gets $60 Million to Spin Off from Beats… But What About That Name?" *Evolver.fm*. http://evolver.fm/2013/03/06/project-daisy-gets-60-million-to-spin-off-from-beats-but-what-about-that-name/

van Couvering, Elizabeth. 2007. "Is Relevance Relevant? Market, Science, and War: Discourses of Search Engine Quality" *Journal of Computer-Mediated Communication* 12(3):866–887.

Van Deusen, Edmund. 1969. "Volney Stefflre: His Bag is People." *Orange Country Illustrated* 7(8):30–33.

Veblen, Thorstein. 1899. *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions*. New York: Macmillan.

Wagner, Roy. 2001. *An Anthropology of the Subject: Holographic Worldview in New Guinea and Its Meaning and Significance for the World of Anthropology*. Berkeley: University of California Press.

Walker, Rob. 2011. "Stealth Iconography: The Waveform." *Design Observer*. http://designobserver.com/feature/stealth-iconography-the-waveform/30008

Watson, Sara. 2014. "Data Doppelgängers and the Uncanny Valley of Personalization." *The Atlantic*. http://www.theatlantic.com/technology/archive/2014/06/data-doppelgangers-and-the-uncanny-valley-of-personalization/372780/

Watson, Sara. 2015. "Data is the New '___.'" *DIS Magazine*. http://dismagazine.com/discussion/73298/sara-m-watson-metaphors-of-big-data/

Weston, Kath. 2008. "'Real Anthropology' and other Nostalgias." In *Ethnographica Moralia: Experiments in Interpretive Anthropology*, edited by Neni Panourgiá and George Marcus. New York: Fordham University Press, 126–137.

Whitman, Brian. "How music recommendation works – and doesn't work." *Variogr.am*. http://notes.variogr.am/post/37675885491/how-music-recommendation-works-and-doesnt-work

Whittaker, Steve and Candace Sidner. 1996. "Email Overload: Exploring Personal Information Management of Email." *CHI '96: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 276–283.

Wieseltier, Leon. 2015. "Among the Disrupted." *New York Times*, January 7. http://www.nytimes.com/2015/01/18/books/review/among-the-disrupted.html

Wikan, Unni. 1992. *Resonance: Beyond the Words*. University of Chicago Press.

Wilf, Eitan. 2013. "Toward an Anthropology of Computer-Mediated, Algorithmic Forms of Sociality." *Current Anthropology* 54(6):716–39.

Willerslev, Rane. 2007. *Soul Hunters: Hunting, Animism, and Personhood among the Siberian Yukaghirs*. Berkeley: University of California Press.

Williams, Brett. 2005. *Debt for Sale: A Social History of the Credit Trap*. Philadelphia: University of Pennsylvania Press.

Williams, Raymond. 1977. *Marxism and Literature*. London: Oxford University Press.

Wilson, Carl. 2011. "The Music Club, 2011." *Slate*. http://www.slate.com/articles/arts/the_music_club/features/2011/music_club_2011/best_music_2011_bon_iver_is_unlistenable_.html

Winner, Langdon. 1986. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago: University of Chicago Press.

Wonder, Stevie. 1976. "Sir Duke." *Songs in the Key of Life*. Motown Records.

Woolgar, Steve. 1990. "Configuring the user: the case of usability trials." In *A Sociology of Monsters: Essays on Power, Technology and Domination*, edited by John Law. Sociological Review Monograph Series.

Wortham, Jenna. 2012. "Taking a Chance on Love, and Algorithms." *New York Times,* April 7. http://www.nytimes.com/2012/04/08/technology/in-online-dating-taking-a-chance-on-love-and-algorithms.html

Zhan, Mei. 2009. *Other-Worldly*. Duke University Press.