Recommended for you: The Netflix Prize and the production of algorithmic culture

Blake Hallinan and Ted Striphas
Indiana University, USA

Abstract
How does algorithmic information processing affect the meaning of the word *culture*, and, by extension, cultural practice? We address this question by focusing on the Netflix Prize (2006–2009), a contest offering US$1m to the first individual or team to boost the accuracy of the company's existing movie recommendation system by 10%. Although billed as a technical challenge intended for engineers, we argue that the Netflix Prize was equally an effort to reinterpret the meaning of *culture* in ways that overlapped with, but also diverged in important respects from, the three dominant senses of the term assayed by Raymond Williams. Thus, this essay explores the conceptual and semantic work required to render algorithmic information processing systems legible as forms of cultural decision making. It also then represents an effort to add depth and dimension to the concept of "algorithmic culture."

"Connect[ing] people to movies they love" is the Netflix company mantra (Netflix Prize, n.d.). It frames the firm as a cheery mediator of people and movies, one that produces delight (and, of course, profit) by fusing technology and subscriber information in a complex alchemy of audiovisual matchmaking. While not inaccurate, such a move mystifies the semantic and socio-technical processes by which these connections are made,

Corresponding author:
Blake Hallinan, Department of Communication & Culture, Indiana University, Classroom Office Building, 800 E. 3rd Street, Bloomington, IN 47405, USA.
Email: bhalli@umail.iu.edu

rendering the Netflix recommendation system a black box, or a kind of known unknown (cf. Madrigal, 2014). Comparable recommendation systems belonging to Amazon, Facebook, Google, Match.com, Microsoft, Twitter, and other technology-driven companies tend to operate similarly, their inner workings "wired shut" with patent and trade secret laws, non-disclosure agreements, non-compete clauses, and other legal instruments (Bottando, 2012; Gillespie, 2007, 2011; Vaidhyanathan, 2011). Although these firms may offer glimpses into their proprietary systems ("Facts about Google and Competition," n.d.; Levy, 2010, 2011), rarely do they invite public conversation about how their algorithms make decisions. In the case of Netflix, however, the recommendation system has not always been so shielded from public scrutiny.

On 6 October 2006, Netflix, Inc., launched the Netflix Prize, a contest offering US$1m to the first individual or team to develop a recommendation system capable of predicting movie ratings with at least 10% greater accuracy than Cinematch, the company's existing system.[1] The competition drew more than 50,000 participants from 186 countries, who organized themselves into roughly 40,000 teams (Netflix Prize Leaderboard, n.d.). Along the way, it prompted more than 9000 posts on the official Netflix Prize discussion forum, where 10,000-plus registered users commiserated about progress and problems relating to the challenge (Netflix Prize Forum, n.d.). Many contestants also utilized personal blogs to chronicle their successes and setbacks, including, notably, members of the championship BellKor's Pragmatic Chaos (BPC) team, and several other contenders who finished in the top twenty. Major news outlets including *BusinessWeek, The New York Times, Slate*, and *Wired* ran features on the event as well. The Netflix Prize concluded on 21 September 2009, just shy of 3 years into the competition and with a nail-biter of an ending. The company received the winning submission just 20 minutes ahead of the entry from the runner-up team, The Ensemble, whose algorithm was on par with BPC's (Lohr, 2009).[2]

We mention all this as a way of underscoring the unusual volume of discourse surrounding Netflix's recommendation algorithm. The competition offers a unique opportunity to scrutinize the details of a system whose peers were, and remain, mostly hidden from view. Furthermore, in the quest to "connect people to the movies they love," the Netflix Prize connected algorithms to art and, in doing so, intervened in the conceptual foundations of culture. We thus reject the nominalist temptation to take the facticity of the competition, qua competition, for granted. Although ostensibly an engineering challenge, the contest was rife with "discourse in which culture addresses its own generality and conditions of existence" (Mulhern, 2000: xiv). As such, we prefer to think about the competition, the contestants, their algorithms, and public statements as parts of a more abstract "situation," a term we borrow from Lauren Berlant to describe a "state of things in which something that will perhaps matter is unfolding amid the usual activity of life." A situation, she continues, "is a state of animated and animating suspension," one "that forces itself on consciousness, that produces a sense of the emergence of something in the present" (Berlant, 2011: 5). The Netflix Prize lends insight, we believe, into how new meanings and practices can insinuate themselves into long-established routines, transforming the latter in ways that may be just reaching popular awareness.

What is the difference, if any, between a human being's determining "the best which has been thought and said," to recall Matthew Arnold's (1993 [1869]: 190) contentious

definition of *culture*, and a computer system's selecting movies tailored to an individual's taste preferences? Or rather, what does *culture* mean, and what might it be coming to mean, given the growing presence of algorithmic recommendation systems such as the one at the center of the Netflix Prize?

These are difficult questions because *culture*, as Raymond Williams (1983) observed, is "one of the two or three most complicated words in the English language" (p. 87). More than 60 years ago, A. L. Kroeber and Clyde Kluckhohn (1963 [1952]) identified 164 different definitions of the word (p. 291). Clearly, it is difficult enough to determine what *culture* means now, let alone what it may be coming to mean in the near future. Williams' (1983) *Keywords* proves particularly helpful in this regard, as he provides an overview of the dominant registers within which operate specific definitions of culture:

> (i) the independent and abstract noun which describes a general process of intellectual, spiritual and aesthetic development, from [the 18th century]; (ii) the independent noun, whether used generally or specifically, which indicates a particular way of life, whether of a people, a period, a group, or humanity in general … But we have also to recognize (iii) the independent and abstract noun which describes the works and practices of intellectual and especially artistic activity. This seems often now the most widespread use: culture is music, literature, painting and sculpture, theatre and film. (p. 90; emphasis removed; see also Williams, 2001 [1961]: 57)

In short, culture may refer to particular modes of fostering human refinement, and their underlying frameworks of valuation and authority; patterns of social difference, commonality, and interaction; and the artifacts, particularly aesthetic objects, associated with specific groups of people.

Together, these rubrics establish a baseline against which to consider burgeoning shifts in *culture*'s definition. The accumulated discourse of the Netflix Prize points to changes in all three registers, demonstrating how "the organization of received meanings has to be made compatible with possible new meanings that are emerging" (Williams, 2001 [1961]: 48). The competition exemplifies how talk about culture has come to occupy the tech world during the last 20 years or so. As we intend to show, engineers now speak with unprecedented authority on the subject, suffusing *culture* with assumptions, agendas, and understandings consistent with their disciplines. The shifting locus of cultural discourse has helped a broad new sense of the word to emerge—one that may be functionally prevalent, we contend, yet vaguely defined. We refer to it as "algorithmic culture": provisionally, the use of computational processes to sort, classify, and hierarchize people, places, objects, and ideas, and also the habits of thought, conduct, and expression that arise in relationship to those processes (Striphas, 2012; see also Galloway, 2006).

This essay considers "how algorithms shape our world" (Slavin, 2011), with particular attention to the conceptual and semantic work required to render algorithmic information processing systems legible as forms of cultural decision making. Our goal is to add depth and dimension to our working definition of algorithmic culture, as well as to make sense of the latter in relationship to long-established frameworks of cultural practice. In doing so, we focus primarily on the production of algorithmic culture. We argue that although the Netflix Prize may have been billed as a competition to improve the company's recommendation algorithm, it was equally an effort to reinterpret what culture is—how it is evaluated, by whom, and to what ends.

Regimes of value

The Netflix Prize attracted machine learning specialists from computer science and, to a lesser extent, mathematicians (Bell et al., 2010: 24). Little wonder, since Netflix, Inc., pitched the competition as a quantitative, technical challenge, typified by the 10% improvement benchmark. That it took almost 3 years to complete, with some contestants claiming to have worked 10–20 hours or more per week on their algorithms, underscores the magnitude of the endeavor (Thompson, 2008). But the participants also discovered the degree to which the technical challenge went hand-in-glove with an interpretive one. Determining the significance of Netflix's customer ratings—along with the broader problem of how to adjudicate cultural values—proved at least as vexing as the engineering.

Netflix invites customers to rate items in its catalog on a scale from 1 to 5 stars, the key to which is revealed as the customer hovers over each star on the website: five stars, "Loved It"; four stars, "Really Liked It"; three stars, "Liked It"; two stars, "Didn't Like It"; and one star, "Hated It." A sixth option, "Not Interested," appears in a separate box below the stars. The system may seem straightforward, yet it has proven a source of consternation for users. In a post on the Netflix company blog, employee "Rubin" (2007) reports having fielded complaints about the system from "a gajillion people." Among the more troubling aspects is the inability to give a film or television program an "average" rating of 2.5, or to render distinctions on the basis of half-stars more generally. Rubin's (2007) colleague, "Todd," followed up in the same blog post to explain how the company interprets ratings:

> For those of you who are concerned that there is no good middle … please consider using 3 stars for that purpose because *that is the way we use it in our recommendation system*. 4–5 star ratings tell us to boost up movies like it when predicting for you and 1–2 star ratings tell us to punish movies like it when we predict for you, but we treat three stars as a neutral signal. (emphasis in original)

While Todd provides clear instruction on the Netflix rating system, his advice only appears on the company blog, a location ancillary to the main website and unlikely to be accessed by a majority of Netflix users. Additionally, his statement on the recommendation system's usage of the ratings belies the more expansive field of implicit context the system has available. Where an editorial movie review offers explicit context for a numerical rating ("four out of five stars!") in the form of written argumentation and implicit context in the reputation of the source and the critic, the Netflix rating system reduces the opportunities for explicit context but continues to draw on implicit contextual information in the form of previous ratings and, more recently, other user data such as location and device. Even if Todd's account is accurate, it is necessarily partial.

The ambiguities in the ratings system flowed downstream to the Netflix Prize competitors, prompting all sorts of questions about the dataset. Especially puzzling were the outliers—the individuals who had rated an unusual amount of movies, or those who had shown a propensity to rate at the extremes. One user, #305344, was guilty of both offenses, having graded in excess of 17,600 movies, 90% of which she or he had rated an average of two points lower compared to other users. In October 2006, contestant

"prodigious" (2006) weighed in on the Netflix Prize discussion forum, suggesting the ratings had to have been "bogus." "Sigmoid Curve" added, "I find it difficult to imagine how rational each and every response to 17000 stimuli could be." In an attempt to interpret the motivation behind user #305344's rations, "FoCu Programmers" replied with the following hypothesis:

> If I am going through the netflix rating system and see some teenie bopper chic flic [*sic*] straight from the Disney channel, I will rate it a 2, maybe even a 1. I don't have to see it to know I won't like it.

Other competitors calculated the amount of time and money it would have taken to watch over 17,000 films, as if to suggest the impossibility of the feat. Adopting a less dismissive approach, competitor "Dishdy" argued, "your algorithm either has to attack these odd-balls or your algorithm has to improve on … your [other] predictions to compensate for those odd-balls" ("A Single Customer That Rated 17,000 Movies," 2006–2007). The contestants may have disagreed about how best to account for customer #305344's evaluations, but all seemed to agree on one thing: she or he was noisy—in the familiar sense of having been something of a ratings loud-mouth, and also in the information-theoretical sense of having been a catalyst for entropy in the dataset (see, for example, Bored Bitless, "A Single Customer That Rated 17,000 Movies," 2006–2007).

Just as Netflix's Todd articulated a proper use of the Netflix rating system on the company blog, these and other contestants struggled to make sense of user behavior they perceived as deviant, or outside of a specific, rational norm. The conversation between Sigmoid Curve, FoCu Programmers, and Dishdy demonstrates a shared imperative: either make the ratings make sense or remove them entirely. The discussion, or rather, the framing of the contest, renders non-rational rating practices superfluous through the presumption that such behaviors either do not exist (with exceptions treated as flawed rationality or corrupt data) or that they are not statistically significant.

One can see a further "noisy" example in *Napoleon Dynamite*, a boutique coming-of-age film released in 2004 by Fox Searchlight. In November 2008, the *New York Times Magazine* ran an extensive article on the Netflix Prize, focusing on the competitors who were then atop the leaderboard and the engineering challenges impeding their progress. Among those challenges was what the *Times* called "the *Napoleon Dynamite* problem." A "very weird and very polarizing" film, *Napoleon Dynamite* drew either intense praise (5 stars) or intense scorn (1 star), but little else, from the Netflix customers who had rated it. It was also frequently an object of disagreement among those whose ratings otherwise aligned closely with one another. Because it was difficult to predict whether users would adore or despise *Napoleon Dynamite*, it skewed the error rate of the competitors' test algorithms to an uncommon degree (Thompson, 2008).

The answer to the *Napoleon Dynamite* problem was less a solution than a work around. The winning entry reckoned the cultural value of particular titles not in absolute but relative terms, situating each rating alongside other ratings the user may have made around the same time and then adjusting the number accordingly. Statistically speaking, a 3-star rating (S_3) might "mean" the same thing as a 4-star (S_4) rating if, in the case of

S_3, a user had just judged five movies worthy of 2 stars, and if, in the case of S_4, a user had just judged a comparable number of titles deserving of 3 stars. A simple average would then give both S_3 and S_4 a "true" rating of 3.5 stars, thus compensating for what contestant Gavin Potter, who finished the competition in seventeenth place, characterized as ratings "inertia"—an apparent tendency among Netflix members to use temporally proximate ratings as a context or "anchor" for subsequent ones (Ellenberg, 2008; Slee, 2009). Here the effect was to attenuate strong opinions by compelling them, algorithmically, to relativize themselves.

Outliers such as "17k guy" and *Napoleon Dynamite* are of interest, then, because they show how contestants debated about cultural artifacts, and, more to the point, how their doing so diverged from Williams' synoptic take on cultural value. While some contestants began by arguing over the merits of a given artifact, most quickly realized the debate did not accomplish much from an information-theoretical standpoint. The latter is not interested in the status of a single user or title but instead in a pattern or system capable of dealing with substantial groups of fungible users and objects. This is why the contestants (or some of them) seemed satisfied with adjusting their variables, so that "problem" reviewers and "problem" films no longer detracted from the predictive power of their algorithms. That is, they had exhausted the value of particular Netflix customers, and of particular titles appearing in its catalog.

All this bespeaks a gradual shift away from debates about great works, or defining canons, to something like the opposite: how to moderate elements of the cultural field that may present themselves as atypical or outstanding, so that they can be led to make sense relative to other, more even-keeled, examples. There may be "no way out of the game of culture," as Pierre Bourdieu (1984: 12) once put it, but what if the rules of the game—and thus the game itself—are changing with respect to algorithmic decision making? With the Netflix Prize, it appears as though questions of cultural authority are being displaced significantly into the realm of technique and engineering, where individuals with no obvious connection to a particular facet of the cultural field (i.e. media) are developing frameworks with which to reconcile those difficult questions. Moreover, issues of quality or hierarchy get transposed into matters of fit, a move that points to how culture is being reimagined as more of a sedentary locus than a trajectory of human development (cf. Pariser, 2011: 130–131).

This shift is explicit in Netflix's corporate pitches, which emphasize an unending circle of customer satisfaction:

> Our passion for helping connect people to great movies is not all altruistic. When people love the movies they watch, they become more passionate about movies, and that helps our business. And as we continue to grow, we are, in turn, able to deliver more movie titles thus increasing customers' delight. As a Netflix employee you can take pride in knowing you are contributing to a service which is so highly enjoyed. (Netflix quoted in Chander, 2008).[3]

The production of sophisticated recommendations produces greater customer satisfaction which produces more customer data which in turn produce more sophisticated recommendations, and so on, resulting—theoretically—in a closed commercial loop in which culture conforms to, more than it confronts, its users.

Ways of life

The Netflix Prize incentivized research about movies and television shows but also about people, suggesting new models of cultural identity latent in the dataset and, presumably, the social. In doing so, the contestants tended to reject dominant (we are tempted to say, "modern") demographic categories in favor of emergent frameworks of identification.

The work of team PragmaticTheory (a forerunner of the championship BPC team) is indicative of the argument that sex, age, race, and other broad classifications fail to capture the more subtle factors relevant to decisions people make about the cultural goods they will consume. "Movie or user data is just not helpful," they state, "because the different algorithms are just too good at capturing the details and nuances that influence user ratings" (*PragmaticTheory*, 2008). *The New York Times Magazine* echoed the point,

> most of the leading teams say that personal information is not very useful, because it's too crude. As one team pointed out to me, the fact that I'm a 40-year-old West Village resident is not very predictive. There's little reason to think the other 40-year-old men on my block enjoy the same movies as I do. (Thompson, 2008)

The assertion flies in the face of the conventional wisdom of marketing, where categories such as age, race, ethnicity, class, gender, and sexuality have long guided "captains of consciousness" in their efforts to pitch products to particular populations (Ewen, 2001; see also Cohen, 2003; Cross, 2000; Curtis, 2002; Zukin, 2004; cf. Duhigg, 2012).[4] Yet the assertion squares with the techno-culture tradition that "conceives of human organisms as nodes in technological networks rather than in social ones," and speaks to the positioning of the contest within the broader history of white masculine engineering culture (Dinerstein, 2006: 585; see also Ensmenger, 2010: 236–237; Marvin, 1988).[5]

On what basis did the contestants construct their algorithms, then? After the conceptual clearing of demographic categories, many came to rely on singular value decomposition or SVD, a technique developed by mathematicians in the field of linear algebra. It refers to a set of procedures for simplifying datasets, on the one hand, and for identifying key terms and internal dependencies, on the other. SVD analysis starts by representing data in the form of a matrix, or an array of values. For instance, the following dataset (which we have created for purposes of illustration) …

ID #	Title	Joy Luck Club	Truman Show	Antonia's Line	Gigli	…	Big Fish
1045				5	1		
3062		4			2		4
0085		4	4				5
8042				5			4
…							
4016		4	3		1		3

… would look like this …

$$M = \begin{bmatrix} 0 & 0 & 5 & 1 & \cdots & 0 \\ 4 & 0 & 0 & 2 & \cdots & 4 \\ 4 & 4 & 0 & 0 & \cdots & 5 \\ 0 & 0 & 5 & 0 & \cdots & 4 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 4 & 3 & 0 & 1 & \cdots & 3 \end{bmatrix}$$

... where each number represents a customer's rating of a given title, and zero functions as a null value in the event of a non-rating. SVD then allows one to reduce the matrix to the product of simpler matrices through a process of factoring, or of breaking down the matrix into more basic components.

Given the unwieldy mathematics of matrix factorization, we will not subject our example to SVD. More important to grasp is the underlying principle, factoring, which is a tool for finding one's way around mathematical expressions. Consider a simpler example: 15×30. There are several ways to factor it: $(3 \times 5) \times (3 \times 10)$; or $(3 \times 5) \times (3 \times [2 \times 5])$; or $3^2 \times (5 \times [2 \times 5])$; or $3^2 \times 5^2 \times 2$. Each factorization reveals something about the original expression. Notice the preponderance of 3s and 5s in the expression $(3 \times 5) \times (3 \times [2 \times 5])$. Factoring allows one to recognize and derive sense from their frequency. Because 3 and 5 are more prevalent, they are weightier than the lone factor, 2. Factoring also provides resources for reducing redundancy through a kind of data compression, as in the prime factorization using exponential notation, $3^2 \times 5^2 \times 2$. SVD functions similarly, allowing one to "format" matrices efficiently and then to determine their significant attributes, or singular values (Austin, n.d.). As Simon Funk (née Brandyn Webb), who is credited with having introduced the Netflix Prize community to SVD, put it,

> Consider just a single column-vector A and corresponding row-vector B. If you multiply A by B you get a matrix as tall as A and as wide as B. Now if you have some target data matrix of the same size (say the Netflix movie-by-user ratings matrix) you can ask: What choice of A and B would make that reconstructed matrix as close to my target matrix as possible? SVD is a mathematical trick for finding that optimal AB pair. (quoted in Piatetsky-Shapiro, 2007)[6]

The word "optimal" is key here, because SVD does not arrive at definitive solutions but rather statements of fit that closely approximate relationships among salient data points.

The turn to SVD was the first "breakthrough" in the Netflix Prize and, arguably, the most significant one (Amatriain, 2009; Thompson, 2008). Although most competitors ended up incorporating other techniques, for many, SVD formed the centerpiece of their algorithms (see, for example, Koren and Bell, 2007; Koren et al., 2009). Its analytical power promised the ability to "teas[e] out very subtle human behaviors" from the dataset and pinpoint previously unidentified cultural categories—"people who like action movies, but only if there's [sic] a lot of explosions, and not if there's a lot of blood. And maybe they don't like profanity," for example (Volinsky, quoted

in Thompson, 2008; cf. Madrigal, 2014). Some contestants even suggested that SVD would allow them to hone in on aspects of reality beyond human perception, language, and sense making:

> For movies, a factor might reflect the amount of violence, drama vs. comedy, more subtle characteristics such as satire or irony, or *possibly some noninterpretable dimension*. For each user, there is a vector in the same dimensional space, with weights that reflect the user's taste for items that score high on the corresponding factor. (Bell et al., 2010: 26, emphasis added; see also Ellenberg, 2008)

This statement would seem to put a new twist on Edward T. Hall's (1966) description of culture as a "hidden dimension," in that whole facets of the cultural realm seem to be unavailable to ethnographers, critics—just about anyone who uses qualitative methods for figuring patterns of social difference, commonality, and interaction.

The Netflix Prize competition was set up to prompt this type of discovery, or at least to make it difficult for the contestants to rely on dominant frameworks of belonging. The dataset included randomly-generated user identification numbers in place of names and, as Netflix legal notes,

> to prevent certain inferences being drawn about the Netflix customer base, some of the rating data for some customers … [has] been deliberately perturbed in one or more of the following ways: deleting ratings; inserting alternative ratings and dates; and modifying rating dates. (Netflix Prize Rules, n.d.)

The dataset listed no further information about the customers; use of information beyond what Netflix had provided constituted a breach of the contest rules. Still, none of this prevented the curious from extracting implicit user information.

Two weeks into the competition, computer scientists Arvind Narayanan and Vitaly Shmatikov posted a draft paper online, explaining how to de-anonymize the dataset. They argued that "removing the identifying information from the records is not sufficient for anonymity" (Narayanan and Shmatikov, 2006–2008: 11) and demonstrated how, by cross-referencing the data with public reviews posted to the Internet Movie Database, they could determine the names of many purportedly anonymous Netflix subscribers. Narayanan and Shmatikov (2006–2008) also pressed further, suggesting that one could infer more intimate aspects of a user's identity such as political leaning, religious affiliation, sexuality, and possibly even body type (p. 16). Their claims were bolstered by *Jane Doe v. Netflix* (2009), a federal class-action lawsuit filed 3 months after the contest's completion. The lead plaintiff in the case, a closeted lesbian mother living in Ohio, feared being outed as a result of having rated numerous gay and lesbian-themed titles through Netflix. "Were her sexual orientation public knowledge," the suit alleged, "it would negatively affect her ability to pursue her livelihood and support her family and would hinder her and her children's ability to live peaceful lives within [her] community" (*Jane Doe v. Netflix*, 2009: §76). Netflix, Inc., settled the suit in March 2010 on undisclosed terms, around the same time the Federal Trade Commission concluded an investigation stemming from privacy concerns relating to

the Netflix Prize. The settlement also resulted in the cancelation of a planned sequel to the contest (Singel, 2010).

What to make of the contestants' claims of having breached a mysterious new frontier of cultural preference and identity, alongside the persistence of established frameworks of belonging? While it might be tempting to choose between these positions or to try to reconcile them, it may be more productive to sustain rather than defuse the tension. Without diminishing the dangers of outing by algorithm, we want to suggest that the stakes of Narayanan and Shmatikov's research, and of the *Jane Doe* case, exceeded the obvious matter of personal privacy.

Echoing previous descriptions, *Wired* characterized the categories obtained by subjecting the Netflix Prize dataset to SVD as "baroque mathematical combinations … that can't be described in words, only in pages-long lists of numbers" (Ellenberg, 2008). The *New York Times* pointed to the "sort of unsettling, alien quality" of the results as evidence that "[t]he machine may be understanding something about us that we do not understand ourselves" (Thompson, 2008). And contestant "Algorist" (2008) dubbed the output "spooky" given the accuracy of the predictions (cf. Madrigal, 2014). The Matrix, the extraterrestrial, the occult: "follow the metaphor," says George Marcus (1998: 92–93), but where do these otherworldly figures lead? Reification, alienation, and mystification, surely, but such appeals offer little assistance in making sense of the streams of numbers and of all the aliens and ghosts suddenly flitting about human surrounds (cf. Latour, 2005). Indeed, it is this sense in which the results hit almost too close to home that is key. It is as if the human lifeworld has been invaded by strangers who nonetheless seem eerily adept at navigating foreign land.

Friedrich Kittler (2006a, 2006b) and Daniel Heller-Roazen (2011) have gone further than anyone in chasing down these metaphors. Kittler (2006b), for his part, has shown how the view of quantification as ill-fitting with ordinary culture depends on the modern bifurcation of system and lifeworld, where numbers, technology, and instrumental reason are viewed as exogenous to human affairs (pp. 40–43). Sometime around 1800, he argues, the humanities arose as such by divesting itself of *quanta*, favoring instead an interpretivist approach that reduced "the interminably open horizon of human existence" to language and experience (Kittler, 2006b: 40). This corresponded with the emergence of an increasingly anthropocentric conceptualization of culture, one that excluded plowshares and other tools with which the term was once associated etymologically (e.g. *coulter*), not to mention plant life and stables-full of non-human animals (Williams, 1983: 87). Likewise, as Kittler and Heller-Roazen have both shown, numbers and mathematics grew partly out of music, today a purportedly cultural activity whose mysteries of tone and harmony led Pythagoras and other pre-Socratic philosophers to a conception of being as that which could be "measured, counted, and weighed" (Kittler, 2006a: 54; see also Heller-Roazen, 2011: 20). The point, then, is that the participants in the Netflix Prize did not stumble upon a previously undiscovered aspect of culture. Instead, they repatriated tools, numbers, and the non-human to a lifeworld from which they had been exiled long ago.

Thus, when John Cheney-Lippold (2011) speaks of "new algorithmic identities" (p. 165), he is correct about the novelty of the specific operations web analytics firms use for parsing cultural identity online. He is also right in cautioning against the ways in which

algorithmic profiling systems "de-essentialize [identity] from its corporeal and societal forms and determinations while … also re-essentializ[ing] [identity] as a statistically-related, largely market research-driven category" (Cheney-Lippold, 2011: 170). Yet, due to the proprietary nature of the systems he studied, Cheney-Lippold, much like Oscar Gandy (1993, 1995) and Greg Elmer (2004) before him, cannot fully adduce how algorithmic identity production exceeds race, class, gender, sexuality, and other familiar categories of cultural identity. Their focus remains on its manifest dimensions, that is on those that are manifest *to people*. As the Netflix Prize competition demonstrates, however, the parameters of human cultural identity stretch beyond the human, all too human to include "prepersonal" or "incorporeal" aspects perceptible to machines (Guattari, 1995: 9).[7] These emerging aspects of cultural identity contain profoundly ambivalent potentialities, and their relationship toward existing modes of personal and cultural identity is far from determined. Will the latent categorizations complement or eclipse extant human understandings? Although the question extends beyond the limits of this article, the example of the Netflix Prize attests to the power of algorithms as they produce categories and understandings that unsettle even their creators.

Objects/ideas/aesthetics

The discoveries of the Netflix Prize shifted the dominant approach to recommendation systems from more traditional collaborative filtering to a blend of latent predictive elements. Yet, despite significant corporate investment in the contest, it turns out that Netflix never fully operationalized the winning algorithm. The decision was driven partly by costs associated with using such an onerous algorithm for commercial data processing but, perhaps more importantly, it seems tied to the fact that, about halfway through the competition, Netflix changed (Amatriain and Basilico, 2012).[8]

Launched in 1997, the company spent its first decade interfacing with clients primarily through its website, where customers created queues of videos they wished to see. Netflix would then ship DVDs to customers by mail, slowly working down the queue as clients returned DVDs in pre-paid envelopes. In February of 2007, Netflix shipped its 1 billionth disk (Netflix Delivers 1 Billionth DVD, 2007), a month after the launch of a new video streaming service. Initially, the streaming service was accessible only through a browser-based web portal, but within a year, both the Xbox game console and Roku set-top streaming device boasted the ability to stream Netflix videos. As the on-demand service took off, other leading developers began incorporating it, too. The service reached the Apple ecosystem in the form of an iOS app and as a feature on Apple TV in 2010, less than a year after the conclusion of the Netflix Prize. An Android app landed a year later, joining a host of other Wi-Fi-connected, Netflix streaming-enabled devices. While Netflix has yet to dispense with DVDs, the move to streaming has fundamentally altered how it conceives of its core business (Amatriain and Basilico, 2012; Schonfeld, 2011). As CEO Reed Hastings, responding to the controversy generated by the 2011 decision (since reversed) to split the video-on-demand service off into a separate company called Qwikster, explains, "we realized that streaming and DVD by mail are becoming two quite different businesses, with very different cost structures, different benefits that need to be marketed differently" (Hastings, 2011).

According to Mohammad Sabah (2012), former Principal Data Scientist for Netflix, the company streamed 2 billion hours of content in the fourth quarter of 2011. That figure doubled in the first quarter of 2013 (Vanderbilt, 2013). Although user ratings continue to play a role in Netflix's recommendations, Sabah (2012) maintains "the implicit signal is stronger." By this he means Netflix now has the capability of tracking when users start, stop, rewind, fast forward, and pause videos, in addition to logging the time of day of viewing, the user's location, the device on which the streaming occurred, whether the user watched a program from beginning to end, what if anything she or he watched next, and more. While Sabah affirms that matrix factorization remains an important data analysis tool for Netflix, the scope and scale of the data far exceed that of the Netflix Prize. More to the point, Netflix is now using data to develop original content in addition to recommending pre-existing material to its subscribers.

In 2013, the company moved aggressively into television,[9] inking exclusive production deals for *House of Cards, Hemlock Grove, Arrested Development*, and *Orange Is the New Black*, following the quiet success of *Lilyhammer*, the company's first foray into original programming, released in February 2012. With *House of Cards*, the company outbid competitors Home Box Office and AMC, landing the show for a reported US$100m (Andreeva, 2011). What drew these firms to the show was the allure of high production-values, typified by the involvement of A-list talent including director David Fincher and star Kevin Spacey, both of whom serve as executive producers. But for Netflix, the deal was not simply about the A-list. According to Ted Sarandos, the company's Chief Content Officer,

> … it's the overlaps that really matter. With *House of Cards*, it was identifying not just somebody who saw *The Social Network* or liked David Fincher but trying to figure out what everybody who liked *Benjamin Button, Seven, Fight Club* and *Social Network* have in common. It's that they love David Fincher's style of storytelling … .You look at Kevin Spacey fans, and then you say, "How about people who love political thrillers?" We went back and pulled all the political thrillers people have watched and rated highly. So you've got all these populations, and right where they overlap in the middle is the low-hanging fruit. If we can get the show in front of these people, they will watch it and love it. (quoted in Rose, 2013)

That is to say, Netflix took a factor-based approach to pursuing *House of Cards*, using its algorithms to decompose the property to determine whether an audience might exist for some combination of "David Fincher," his "style," the collection of genres across which he has worked, "Kevin Spacey," the specific genre of political thriller, and so forth. Moreover, this "low-hanging fruit" differs from the more frequent articulation of mass media as appealing to the "lowest common denominator" in the size and specificity of the intended audience, with Netflix moving away from an undifferentiated mass toward an aggregation of highly differentiated micro-audiences.

Data collection and interpretation permeate many aspects of corporate decision making, from the vetting of potential acquisitions to the shaping of the context of acquired properties. Netflix is not simply hiring *auteurs* whose unique vision for a production prevails over all else. In the case of *Orange Is the New Black*, Sarandos reports that Netflix has exerted "a lot of casting influence" over the property based on what its

algorithms suggested would be the most effective choices for actors in terms of attracting audiences and new subscribers (quoted in Rose, 2013). The company has taken an equally radical step in choosing to release an entire season of its shows all at once, rather than doling out one new episode per week at a regularly scheduled time. The shift away from "appointment viewing," long prevalent in traditional television, to "binge viewing" grew out of Netflix's analysis of viewing data, which showed its streaming customers tended to watch several TV episodes back to back instead of one at a time. The insight has affected both the structure and content of these shows, allowing scriptwriters to side-step recaps, cliff-hangers, and similar narrative devices intended to keep viewers glued between commercial breaks and from one week to the next (Rose, 2013). While Netflix has not extensively described the relationship between its viewing data and production decisions, Sarandos has affirmed that the company's goal is "to optimize for the best shows," informed by "data-driven hunches" (quoted in Karpel, 2012).

Of course, data are an enduring aspect of the media industries. Long before the rise of "big data," studio executives turned to test audiences, ratings, viewer diaries, eye track-ing, and unsolicited feedback in an effort to gauge reception of their programming.[10] Industry professionals continue to use these and other resources to guide their decisions about what material to produce in addition to how, and with whom, to populate the con-tent. Yet, there may be something different now emerging with respect to the *addressivity* of culture—a tendency that became apparent, though dimly, with the Netflix Prize and that is becoming more apparent as the company ratchets up its original shows.

Williams' third and final definitional register encompasses a host of artifactual under-standings of *culture*. They refer, at their broadest, to the practical, material, and intellec-tual "stuff" of a people, and, at their most specific, to a subset of objects and activities determined to be aesthetic. This tension between the general and the particular is a cen-tral stake in this rubric, and it has much to do with whether one takes culture to be "ordi-nary" (Williams, 1989), "the best which has been thought and said" (Arnold, 1993 [1869]), or something else in between (see, for example, Macdonald, 1961; Radway, 1997; Rubin, 1992).[11] But Netflix, and indeed the Netflix Prize, suggests more is at stake in this definition than which aspects of a people's expressive life get to count as culture. What matters increasingly is the "universe of reference" implicit in this rubric (Guattari, 1995: 9), or the agency to which artifacts must appeal, as it were, to gain entry to culture. Until recently, one might reasonably surmise this agency was limited to human beings. Because people appeared to determine the worthiness of prospective cultural artifacts, more or less uniquely, it followed that culture's universe of reference must be endoge-nous to the human species. Little wonder that Williams (1958) described *culture* in his early work as a "court of *human* appeal" (p. xviii; emphasis added). As the public com-mentary on production decisions behind *House of Cards* and *Orange Is the New Black* suggest, expressive materials now find themselves addressing computers, too. These sys-tems, and more importantly their algorithms, play a critical role in deciding which arti-cles (or parts thereof) gain admission to the cultural realm, and in what form. Their doing so thus points in the direction of another universe of reference—a court of *algorithmic* appeal in which objects, ideas, and practices are heard, cross-examined, and judged inde-pendently, in part, of human beings.[12]

Conclusion

The case of the Netflix Prize demonstrates the significance of contest design on the development and social consequences of a recommendation system. It also affirms the importance of situating any analysis of algorithmic culture in the details of cultural production. The extensive archive of the Netflix Prize, specifically the official contest forums, the competitors' online presence, and the resulting scholarly literature on the science of recommendation systems, testifies to the amount of labor involved in making an algorithmic culture. At the same time, ideas about cultural judgment, ways of life, and products developed within the context of the Netflix Prize proved difficult to explain to broader audiences because of the degree to which those ideas appear to have run afoul of more established ways of understanding *culture*. The consistent focus of press coverage on the spectacle of the contest, replete with trite descriptions of socially-awkward computer scientists and glorious send-ups of the lone hacker on the range (see, for example, Ellenberg, 2008; Thompson, 2008), elides investigation into the content and its consequences. The contestants' anecdotes about explaining their work to outsiders also demonstrate the difficulty in substantively discussing algorithmic culture beyond a community of experts, journalist Alexis Madrigal's (2014) effort to "reverse engineer" the "grammar" of Netflix's seemingly idiosyncratic genres notwithstanding.

While the semantics of algorithmic culture may struggle with portability, the practices and processes are more easily exported. Some members of Team Ensemble brought their skills to Opera Solutions, a consulting firm providing recommendation system tools to other companies and advertisers, while contestant Yehuda Koren's advances in modeling the temporal dynamics of ratings have helped bring more desired content to Yahoo! news homepages, to only scratch the surface of influence (Krakovsky, 2009; Manjoo, 2009). Netflix, meanwhile, continues to use principles developed during the contest in its recommendation system, which influences upward of 75% of all user selections. And Netflix represents no small share of the streaming video-on-demand market—the source of nearly one-third of all North American downstream internet traffic at peak hours, consuming approximately 50% more bandwidth than its nearest competitor, YouTube (Bump, 2013; Schonfeld, 2011; Spangler, 2013).

More broadly, this essay represents an inquiry into "the unraveling of the associations which had earlier sustained the meaning of culture" (Bennett, 2005: 64). This is not to suggest the abandonment of those associations, but rather to point to the emergence of different meanings whose presence reframes dominant conceptions of the term. The tensions we have traced—from "great works" to more tepid forms, from personal to prepersonal renderings of cultural identity, and from a court of human appeal to an algorithmic one—are not intended as dichotomies. Although there may be little consensus as to what *culture* means, most everyone who has studied the term seems to agree on its elasticity, or its ability to manage multiple, sometimes competing, demands (Williams, 1983: 89–90). The story of the Netflix Prize is not about one set of definitional frameworks superseding another, but about a jockeying for position as dominant understandings of *culture* bump up against emergent ones—something apparent in the contestants' disinterest in established ways of parsing cultural value, for instance, and also in the *Jane Doe* case. It is a story about latencies, moreover, or the reactivation of long-dormant ways of

conceiving of cultural life—for example, in terms of mathematical principles. And it is a story, finally, about tendencies toward the future, a future in which people and algorithms will continue to become ever more entangled in cultural production, both on and offline (Gillespie, 2014: 183–184).

Furthermore, the Netflix Prize both complements and clarifies accounts of "how we became posthuman." Drawing on the history of cybernetics and information theory, N. Katherine Hayles (1999) has explored how the emergence of new classes of computational devices, together with advanced conceptualizations of signal processing, have challenged classically-liberal understandings of human subjectivity, particularly the notion of the integral self and its difference from organized inorganic matter. She is interested in human–machine entanglements and, by extension, how new technological, ideational, and semantic contexts alter what it means to be human. Culture is not a central tenet of the work, although it clearly subtends throughout, notably in the discussions of books and literature in the information age. Although the characters, setting, and approach may differ, our discussion of culture and the Netflix Prize closely parallels what Hayles has to say about the figure of the human as it relates to cybernetic machines. The difficulty lies in the extent of Hayles' claims, which oscillate between the carefully situated—"'human' and 'posthuman' coexist in shifting configurations that vary with historically specific contexts" (Hayles, 1999: 6)—and the transhistorical—"we have always been posthuman" (Hayles, 1999: 291). The case of the Netflix Prize suggests the former is the desirable view, analytically, given the prevalence of residual, dominant, and emergent elements in discourse (Williams, 1977).

The Netflix Prize also raises challenging questions. What happens when engineers—or their algorithms—become important arbiters of culture, much like art, film, and literary critics? How do we contest computationally-intensive forms of identification and discrimination that may be operating in the deep background of people's lives, forms whose underlying mathematical principles far exceed a reasonable degree of technical competency? What is at stake in "optimizing" would-be cultural artifacts to ensure a more favorable reception, both by human audiences and by algorithms? The Netflix Prize opens up these questions, and though it hardly settles them, it nonetheless offers needed perspective on what *culture* may be coming to mean. Indeed, if *culture* is not exactly what it once was, then this is all the more reason to make sense of it anew. Otherwise, we risk hampering our ability to participate meaningfully in a world in which culture and computation are becoming less distinguishable from one another.

Funding

Notes

1. Netflix provided contestants with access to two datasets: a "training set" consisting of approximately 100 million user ratings, dated but anonymized, of 17,700 films/television programs, and a "qualifying set" listing user identification numbers, the items they had rated, and the dates they had rated particular items, but no actual ratings. The goal of the contest was to create an algorithm capable of predicting how users in the qualifying set had rated a particular

item, based on patterns inferred from the training set (Netflix Prize Rules, n.d.).

2. The Ensemble's algorithm was, technically, slightly more accurate than that of BPC but was considered equal to the winner's because of rounding (Miller, 2009).

3. The quote is from a list of reasons to work at the company, originally posted on the Netflix jobs page (https://www.netflix.com/Jobs?id=5366) in 2007, but as of this writing the site is no longer available. Additionally, the page includes code that prevents it from being crawled by Internet Archive and other web-preservation organizations. However, a number of secondary sources confirm the existence and relevance of the page ("Ever Wonder Why Netflix Was Doing So Well?," 2007; McCord, 2014; "Netflix—A Powerful Example in Branding," 2007; Patel, 2007; Schneider, n.d.).

4. Advertisers have also developed ways of parsing the social that are more or less independent of traditional demographic categories—for example, "values and lifestyles," "relationship marketing," and so on (Curtis, 2002; Elmer, 2004: 9–10, 68; Zukin, 2004: 108–109).

5. This is not to claim all contestants were both white and male, but rather to argue that the contest—from the datasets to the participants—operated in a world where race, gender, and other social relationships or categorizations were only as relevant as their predictive power, and even this possibility is largely presumed non-existent. Despite the attempt to insulate the contest, and thus recommendation systems, from the stakes of social relations, the boundary is far from discrete.

    The Netflix Prize Discussion Forum thread "Female Contestants" (2009) offers a notable exception, although the discussion of gender surfaces with regard to the contestants rather than the dataset. The posts that follow the initial prompt range from misogynistic jokes to speculations of cultural differences, to an analysis of structure and socialization, and the unique experiences of gender discrimination as someone who identifies as genderqueer.

6. Singular value decomposition (SVD) works by factoring the original matrix into a product three, rather than two, matrices. Funk seems to be simplifying for explanatory purposes.

7. Hayles' (1999) discussion of machine reading offers a useful parallel with regard to the decomposition of printed books: "the computer will garble its [the book's] body, breaking it apart and reassembling it into the nonstory of a data matrix rather than an entangled and entangling narrative" (p. 41).

8. Netflix implemented an earlier version developed by three members of the winning team, which showed an 8.43% improvement over Cinematch (Amatriain and Basilico, 2012).

9. We use the word "television" advisedly, since the content is not television in the familiar sense of the word. For more on post-television, see Spigel and Olsson (2004), Lotz (2007) and Tryon (2013).

10. To wit, "On the basis of ratings, fortunes rose and fell, and heads rolled with unnerving suddenness. Contempt for the needs of 'the market' was often heard, especially from writers, directors, and actors, but the rhetoric of denunciation seldom changed anything" (Barnouw, 1990: 404; on letters, see Bodroghkozy, 1992: 148).

11. And in this, Williams (1983: 90–91) notes the intimate association between rubrics one (culture as valuation) and three (culture as artifacts).

12. This is akin to the generalization of Search Engine Optimization (SEO) beyond the web. Or, as Gillespie (2014) describes it,

    Just as Hollywood's emphasis on specific genres invites screenwriters to write in generic ways, the Flickr algorithm may induce subtle reorientations of photographers' practices toward its own constructed logic, that is, toward aspiring to photograph in ways adherent to certain emergent categories, or orienting their choice of subject and composition toward those things the algorithm appears to privilege. (pp. 183–184)

References

Algorist (2008) Need account to get recommendations. *Forum Post*. Available at: http://www. netflixprize.com/community/viewtopic.php?pid=6147-p6147 (accessed 19 November 2013).

Amatriain X (2009) The Netflix Prize: lessons learned. *TechnoCalifornia*, 29 September. Available at: http://technocalifornia.blogspot.com/2009/09/netflix-prize-lessons-learned. html (accessed 19 November 2013).

Amatriain X and Basilico J (2012) Netflix recommendations: beyond the 5 stars (part 1). *Netflix Tech Blog*, 6 April. Available at: http://techblog.netflix.com/2012/04/netflix-recommenda-tions-beyond-5-stars.html (accessed 17 June 2013).

Andreeva N (2011) Netflix to enter original programming with mega deal for David Fincher-Kevin Spacey series "House Of Cards." *Deadline.com*, 15 March. Available at: http://www. deadline.com/2011/03/netflix-to-enter-original-programming-with-mega-deal-for-david-fincher-kevin-spacey-drama-series-house-of-cards/ (accessed 19 November 2013).

Arnold M (1993 [1869]) *Culture and Anarchy and Other Writings*. Cambridge; New York: Cambridge University Press.

"A Single Customer That Rated 17,000 Movies" (2006–2007) *Forum*. Available at: http://www. netflixprize.com//community/viewtopic.php?id=141 (accessed 18 November 2013).

Austin D (n.d.) We recommend a singular value decomposition. *Feature Column*. Available at: http://www.ams.org/samplings/feature-column/fcarc-svd (accessed 18 November 2013).

Barnouw E (1990) *Tube of Plenty: The Evolution of American Television*. New York: Oxford University Press.

Bell RM, Koren Y and Volinsky C (2010) All together now: a perspective on the Netflix Prize. *Chance* 23(1): 24–29.

Bennett T (2005) Culture. In: Bennett T, Grossberg L and Morris M (eds) *New Keywords: A Revised Vocabulary of Culture and Society*. Malden, MA: Blackwell, pp. 63–69.

Berlant L (2011) *Cruel Optimism*. Durham, NC: Duke University Press.

Bodroghkozy A (1992) "Is This What You Mean by Color TV?": race, gender, and contested meanings in NBC's *Julia*. In: Spigel L and Mann D (eds) *Private Screenings: Television and the Female Consumer*. Minneapolis, MN: University of Minnesota Press, pp. 143–167.

Bottando E (2012) *Hedging the Commons: Google Books, Libraries, and Open Access to Knowledge*. PhD Dissertation, Department of Communication Studies, University of Iowa, Iowa City, IA.

Bourdieu P (1984) *Distinction: A Social Critique of the Judgment of Taste* (trans. R Nice). Cambridge, MA: Harvard University Press.

Bump P (2013) Half of Internet traffic in North America is just to watch Netflix and YouTube. *The Wire*, 14 May. Available at: http://www.thewire.com/technology/2013/05/netflix-youtube-traffic/65210/ (accessed 19 November 2013).

Chander (2008) More on "A new Corporate Organization Model." *CDOQ*, 30 December. Available at: http://cdoq.blogspot.com/2008/12/more-on-new-corporate-organization.html (accessed 20 March 2014).

Cheney-Lippold J (2011) A new algorithmic identity: soft biopolitics and the modulation of control. *Theory, Culture & Society* 28(6): 164–181.

Cohen L (2003) *A Consumers' Republic: The Politics of Mass Consumption in Postwar America*. New York: Knopf.

Cross GS (2000) *An All-Consuming Century: Why Commercialism Won in Modern America*. New York: Columbia University Press.

Curtis A (2002) *The Century of the Self* (240 minutes). BBC Four.

Dinerstein J (2006) Technology and its discontents: on the verge of the posthuman. *American Quarterly* 58(3): 569–595.

Duhigg C (2012) How companies learn your secrets. *The New York Times*, 16 February. Available at: http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html (accessed 18 November 2013).

Ellenberg J (2008) This psychologist might outsmart the math brains competing for the Netflix Prize. *Wired*, March. Available at: http://archive.wired.com/techbiz/media/magazine/16-03/mf_netflix?currentPage=all (accessed 18 November 2013).

Elmer G (2004) *Profiling Machines: Mapping the Personal Information Economy*. Cambridge, MA: MIT Press.

Ensmenger N (2010) *The Computer Boys Take Over: Computers, Programmers, and the Politics of Technical Expertise*. Cambridge, MA: MIT Press.

"Ever Wonder Why Netflix Was Doing So Well?" (2007) Discussion forum, 17 October. Available at: http://www.reddit.com/comments/5yhx5/every_wonder_why_netflix_is_doing_so_well_check (accessed 2 April 2014).

Ewen S (2001) *Captains of Consciousness: Advertising and the Social Roots of the Consumer Culture*. 25th anniversary ed. New York: Basic Books.

"Facts about Google and Competition" (n.d.) *Google Competition*. Available at: http://googlecompetition.blogspot.com (accessed 6 June 2014).

"Female Contestants" (2009) *Forum*. Available at: http://www.netflixprize.com/community/viewtopic.php?pid=9141 (accessed 1 April 2014).

Galloway AR (2006) *Gaming: Essays on Algorithmic Culture*. Minneapolis, MN: University of Minnesota Press.

Gandy OH (1993) *The Panoptic Sort: A Political Economy of Personal Information*. Boulder, CO: Westview Press.

Gandy OH (1995) It's discrimination, stupid! In: Brook J and Boal I (eds.) *Resisting the Virtual Life: The Culture and Politics of Information*. San Francisco, CA: City Lights Books, pp. 35–47.

Gillespie T (2007) *Wired Shut: Copyright and the Shape of Digital Culture*. Cambridge, MA: MIT Press.

Gillespie T (2011) Can an algorithm be wrong? Twitter trends, the specter of censorship, and our faith in the algorithms around us. *Culture Digitally*. Available at: http://culturedigitally.org/2011/10/can-an-algorithm-be-wrong/ (accessed 18 November 2013).

Gillespie T (2014) The relevance of algorithms. In: Gillespie T, P Boczkowski P and Foot K (eds) *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: MIT Press, pp. 167–194.

Guattari F (1995) *Chaosmosis: An Ethico-Aesthetic Paradigm* (trans. P Bains and J Pefanis). Bloomington, IN: Indiana University Press.

Hall ET (1966) *The Hidden Dimension*. Garden City, NY: Doubleday.

Hastings R (2011) An explanation and some reflections. *Netflix US & Canada Blog*, 18 September. Available at: http://blog.netflix.com/2011/09/explanation-and-some-reflections.html (accessed 10 November 2013).

Hayles K (1999) *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, IL: University of Chicago Press.

Heller-Roazen D (2011) *The Fifth Hammer: Pythagoras and the Disharmony of the World*. New York: Zone Books.

Jane Doe v. Netflix (2009) Original complaint, 17 December (United States District Court for the Northern District of California). Available at: http://www.wired.com/images_blogs/threatlevel/2009/12/doe-v-netflix.pdf (accessed 10 December 2013).

Karpel A (2012) Netflix's head of content Sarandos queues up an original programming strategy. *Fast Company Create*, 2 February. Available at: http://www.fastcocreate.com/1679595/

netflixs-head-of-content-sarandos-queues-up-an-original-programming-strategy (accessed 19 November 2013).

Kittler F (2006a) Number and numeral. *Theory, Culture & Society* 23(7–8): 51–61.

Kittler F (2006b) Thinking colours and/or machines. *Theory, Culture & Society* 23(7–8): 39–50.

Koren Y and Bell R (2007) Lessons from the Netflix Prize Challenge. *SIGKDD Explorations Newsletter* 9(2): 75–79.

Koren Y, Bell R and Volinsky C (2009) Matrix factorization techniques for recommender systems. *Computer* 42(8): 30–37.

Krakovsky M (2009) Award-winning paper reveals key to Netflix Prize. *Communications of the ACM*. Available at: http://cacm.acm.org/news/32450-award-winning-paper-reveals-key-to-netflix-prize/fulltext (accessed 10 December 2013).

Kroeber AL and Kluckhohn C (1963 [1952]) *Culture: A Critical Review of Concepts and Definitions*. New York: Vintage.

Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford: Oxford University Press.

Levy S (2010) How Google's algorithm rules the web. *Wired*, March. Available at: http://www.wired.com/magazine/2010/02/ff_google_algorithm/ (accessed 17 September 2013).

Levy S (2011) *In the Plex: How Google Thinks, Works, and Shapes Our Lives*. New York: Simon & Schuster.

Lohr S (2009) A $1 million research bargain for Netflix, and maybe a model for others. *The New York Times*, 22 September. Available at: http://www.nytimes.com/2009/09/22/technology/internet/22netflix.html (accessed 18 November 2013).

Lotz AD (2007) *The Television Will Be Revolutionized*. New York: New York University Press.

McCord P (2014) How Netflix reinvented HR. *Harvard Business Review*, January–February. Available at: http://hbr.org/2014/01/how-netflix-reinvented-hr/ar/1 (accessed 2 April 2014).

Macdonald D (1961) *Masscult and Midcult* (Partisan Review). New York: NYRB Classics.

Madrigal A (2014) How Netflix reverse engineered Hollywood. *The Atlantic*, 2 January. Available at: http://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679/ (accessed 29 March 2014).

Manjoo F (2009) The Netflix Prize was brilliant. *Slate*. Available at: http://www.slate.com/articles/technology/technology/2009/09/the_netflix_prize_was_brilliant.html (accessed 10 December 2013).

Marcus GE (1998) *Ethnography through Thick and Thin*. Princeton, NJ: Princeton University Press.

Marvin C (1988) *When Old Technologies Were New: Thinking About Electric Communication in the Late Nineteenth Century*. New York: Oxford University Press.

Miller A (2009) Netflix Prize conclusion. *The Ensemble*, 22 September. Available at: http://www.the-ensemble.com/content/netflix-prize-conclusion (accessed 18 November 2013).

Mulhern F (2000) *Culture/Metaculture*. London; New York: Routledge.

Narayanan A and Shmatikov V (2006–2008) How to break anonymity of the Netflix Prize dataset. *arXiv e-print*. Available at: http://arxiv.org/abs/cs/0610105 (accessed 18 November 2013).

Netflix Delivers 1 Billionth DVD (2007) *Msnbc.com*, 25 March. Available at: http://www.nbc-news.com/id/17331123/ns/business-us_business/t/netflix-delivers-billionth-dvd/ (accessed 19 November 2013).

Netflix Prize Forum (n.d.) http://www.netflixprize.com/community/ (accessed 18 November 2013).

Netflix Prize Leaderboard (n.d.) http://www.netflixprize.com/leaderboard (accessed 18 November 2013).

Netflix Prize (n.d.) http://www.netflixprize.com (accessed 18 November 2013).

Netflix Prize Rules (n.d.) http://www.netflixprize.com//rules (accessed 18 November 2013).

"Netflix—A Powerful Example in Branding" (2007) Image and comment, 17 October. Available at: https://www.flickr.com/photos/bruceclay/1606068204/ (accessed 2 April 2014).

Pariser E (2011) *The Filter Bubble: What the Internet Is Hiding From You*. New York: Penguin Press.

Patel N (2007) Learning from Netflix's rules for success. *Search Engine Journal*, 18 October. Available at: http://www.searchenginejournal.com/learning-from-netflixs-rules-for-success/55282/ (accessed 2 April 2014).

Piatetsky-Shapiro G (2007) Interview with Simon Funk: why SVD approach? *KDnuggets*, 24 April. Available at: http://www.kdnuggets.com/news/2007/n08/6i.html (accessed 18 November 2013).

Prodigious (2006) *A single customer that rated 17,000 movies. Forum Post*. Available at: http://www.netflixprize.com/community/viewtopic.php?pid=627#p627 (accessed 6 June 2014).

PragmaticTheory (2008) You want the truth, you can't HANDLE the truth! *PragmaticTheory*, 3 August. Available at: http://pragmatictheory.blogspot.com/2008/08/you-want-truth-you-cant-handle-truth.html (accessed 18 November 2013).

Radway JA (1997) *A Feeling for Books: The Book-of-the-Month Club, Literary Taste, and Middle-Class Desire*. Chapel Hill, NC: University of North Carolina Press.

Rose L (2013) Netflix's Ted Sarandos reveals his "Phase 2" for Hollywood. *The Hollywood Reporter*, 22 May. Available at: http://www.hollywoodreporter.com/news/netflixs-ted-sarandos-reveals-his-526323 (accessed 19 November 2013).

Rubin JS (1992) *The Making of Middlebrow Culture*. Chapel Hill, NC: University of North Carolina Press.

Rubin (2007) Todd speaks: 1/2 Stars. *Netflix Tech Blog*, 6 August. Available at: http://blog.netflix.com/2007/08/todd-speaks-12-stars.html (accessed 18 November 2013).

Sabah M (2012) Mohammad Sabah—Hadoop Summit 2012—theCUBE. Available at: http://www.youtube.com/watch?v=9h6wQXuke_E&feature=youtube_gdata_player (accessed 19 November 2013).

Schneider L (n.d.) Working at Netflix—reasons to work at Netflix. *About.com*. Available at: http://jobsearchtech.about.com/od/companyprofiles/a/Netflix_3.htm (accessed 2 April 2014).

Schonfeld E (2011) Netflix now the largest single source of Internet traffic in North America. *TechCrunch*, 17 May. Available at: http://techcrunch.com/2011/05/17/netflix-largest-internet-traffic/ (accessed 19 November 2013).

Singel R (2010) Netflix cancels recommendation contest after privacy lawsuit. *Wired*, 12 March. Available at: http://www.wired.com/threatlevel/2010/03/netflix-cancels-contest/ (accessed 19 November 2013).

Slavin K (2011) How algorithms shape our world. *TEDGlobal*. Available at: http://www.ted.com/talks/kevin_slavin_how_algorithms_shape_our_world.html (accessed 15 February 2013).

Slee T (2009) Netflix Prize: was the *Napoleon Dynamite* problem solved? Available at: http://whimsley.typepad.com/whimsley/2009/10/netflix-prize-was-the-napoleon-dynamite-problem-solved.html (accessed 18 November 2013).

Spangler T (2013) Netflix video puts even more strain on the Internet. *Variety*, 14 May. Available at: http://variety.com/2013/digital/news/netflix-puts-even-more-strain-on-the-internet-1200480561/ (accessed 19 November 2013).

Spigel L and Olsson J (eds) (2004) *Television after TV: Essays on a Medium in Transition*. Durham, NC: Duke University Press.

Striphas T (2012) What is an Algorithm? *Culture Digitally*. Available at: http://culturedigitally.org/2012/02/what-is-an-algorithm/ (accessed 15 February 2013).

Thompson C (2008) If you liked this, you're sure to love that. *The New York Times*, 23 November. Available at: http://www.nytimes.com/2008/11/23/magazine/23Netflix-t.html (accessed 18 November 2013).

Tryon C (2013) *On-Demand Culture: Digital Delivery and the Future of Movies*. New Brunswick, NJ: Rutgers University Press.

Vaidhyanathan S (2011) *The Googlization of Everything (and Why We Should Worry)*. Berkeley, CA: University of California Press.

Vanderbilt T (2013) The science behind the Netflix algorithms that decide what you'll watch next. *Wired*, August. Available at: http://www.wired.com/underwire/2013/08/qq_netflix-algorithm/ (accessed 19 November 2013).

Williams R (1958) *Culture and Society, 1780–1950*. New York: Columbia University Press.

Williams R (1977) *Marxism and Literature*. Oxford: Oxford University Press.

Williams R (1983) *Keywords: A Vocabulary of Culture and Society*. Rev. ed. New York: Oxford University Press.

Williams R (1989) Culture is Ordinary. In: *Resources of Hope: Culture, Democracy, Socialism*, London: Verso, pp. 3–18.

Williams R (2001 [1961]) *The Long Revolution*. Orchard Park, NY: Broadview Press.

Zukin S (2004) *Point of Purchase: How Shopping Changed American Culture*. New York: Routledge.

Author biographies

Blake Hallinan is a PhD student in the Department of Communication and Culture at Indiana University. Her research focuses on the social and technical determinations of preference within media industries. Email: bhalli@indiana.edu; Twitter:@blakeplease.

Ted Striphas is Associate Professor in the Department of Communication and Culture, Indiana University. He is author of *The Late Age of Print: Everyday Book Culture from Consumerism to Control* (Columbia U.P., 2009) and is currently at work on his next book, *Algorithmic Culture.* Email: striphas@indiana.edu; Twitter:@striphas.