"Raw Data" Is an Oxymoron

۲

Edited by Lisa Gitelman

The MIT Press Cambridge, Massachusetts London, England

Gitelman—"Raw Data" Is an Oxymoron

۲

۲

© 2013 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

۲

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Perpetua by Toppan Best-set Premedia Limited, Hong Kong. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

••

۲

10 9 8 7 6 5 4 3 2 1

Contents

Acknowledgments vii

Introduction 1 Lisa Gitelman and Virginia Jackson

- 1 Data before the Fact 15 Daniel Rosenberg
- 2 Procrustean Marxism and Subjective Rigor: Early Modern Arithmetic and Its Readers 41 Travis D.Williams

۲

- From Measuring Desire to Quantifying Expectations: A Late Nineteenth-Century Effort to Marry Economic Theory and Data 61
 Kevin R. Brine and Mary Poovey
- Where Is That Moon, Anyway?: The Problem of Interpreting Historical Solar Eclipse Observations 77 Matthew Stanley
- 5 "<u>facts</u> and FACTS": Abolitionists' Database Innovations 89 Ellen Gruber Garvey
- 6 Paper as Passion: Niklas Luhmann and His Card Index 103
 Markus Krajewski

۲

()

vi Contents

- 7 Dataveillance and Countervailance 121 Rita Raley
- 8 Data Bite Man: The Work of Sustaining a Long-Term Study 147 David Ribes and Steven J. Jackson

۲

Data Flakes: An Afterword to Raw Data Is an Oxymoron 167 Geoffrey C. Bowker

List of Contributors 173

Index 179

۲

7 Dataveillance and Countervailance

Rita Raley

It's what we call a massive data-base tally. Gladney, J.A.K. I punch in the name, the substance, the exposure time and then I tap into your computer history. Your genetics, your personals, your medicals, your psychologicals, your police-and-hospitals. It comes back pulsing stars. This doesn't mean anything is going to happen to you as such, at least not today or tomorrow. It just means you are the sum total of your data. No man escapes that.

—Don DeLillo, White Noise

What is most unfortunate about this development is that the data body not only claims to have ontological privilege, but actually has it. What your data body says about you is more real than what you say about yourself. The data body is the body by which you are judged in society, and the body which dictates your status in the world. What we are witnessing at this point in time is the triumph of representation over being.

-Critical Art Ensemble

The Data Bubble

As I set about the process of wiping my machine of all cookies a few summers ago in preparation for the cloning of my hard drive, I was somewhat naively surprised to learn about so-termed Flash cookies, or LSOs (local shared objects). Internet privacy has always been a concern: I have long made a point of systematically deleting cookies along with my cache and search history, researching the plug-ins and extensions best able to anonymize my browsing, and using search engines that do not record IP addresses, particularly those that work against search leakage.¹ I have also made a point of providing false personal information and developing a suite of pseudonymous identities (user names, avatars, anonymous email addresses), the purpose of which has been to convince myself that I am able to maintain some aspect of control over my own data. My error

R

was thinking within the architecture of the browser window: LSOs are, as the name suggests, local cookies stored outside of the browser, in my case at rraley/Library/ Preferences/Macromedia/Flash Player/#SharedObjects, and thus not deletable from a browser toolbar.² In basic terms, LSOs are tracking devices within the Flash player that override the user's security preferences and are set without her knowledge and consent. There are applications such as Flush and BetterPrivacy that will ostensibly manage and clean out LSOs, but their most pernicious aspect is their capacity to "respawn" tracking cookies with data stored in Flash; that is, Flash Local Storage is used to back up the HTML cookies for the explicit purpose of restoring them within seconds after they are deleted. These zombie cookies-and this is certainly their effect, as manual deletion and even the aforementioned tools are essentially futile—are made possible by what Adobe Systems insists is a "misuse" of Local Storage, though it is worth noting that the privacy settings panel on Adobe's site is notoriously difficult to read, appearing as a demo rather than as an actual window.³ They have not then been invisible to me alone, though the larger issue of data collection continues to receive more public attention in the wake of investigative reports such as the Wall Street Journal's "What They Know" series.⁴

The immediate purpose of LSOs, along with traditional and third-party cookies, is online behavioral advertising, the economic potential of which is no doubt clear: consider the speculative value of the uniquely numbered cookie assigned to each machine, one that collates ostensibly nonpersonal behavioral information in order to produce a closely approximate demographic portrait including age, gender, location, educational level, income, consumption habits (purchasing and reading), sexual preference, and health issues.5 The "audience management experts" of Demdex, Inc., for example, transform the profile of a common user into one of a unique individual by combining the ID code from a single machine, one that holds a summary record of browsing and search history, with offline data including census information, real estate records, and car registration.° As John Battelle puts it, this information is producing "a massive clickstream database of desires, needs, wants, and preferences that can be discovered, subpoenaed, archived, tracked, and exploited for all sorts of ends."7 Online behavioral advertising produces a dynamic, flexible, and perfectly customized audience, constituted by the microtargeting of the intents and interests of consumers on a massive scale. In practical terms, if a consumer happens upon but fails to make a purchase from a particular retail site that aligns with her profile, that microtargeting can become retargeting, which means that ads for an item she has viewed will be pushed to other nonretail sites, or to adopt the rhetoric of personalized retargeting companies, she will be found as she browses and driven back to the original site. In its ultimate form, such a targeting system would locate a user in close proximity to a shopping market, assess the whole of her shopping history, compare those purchases with those of other shoppers, and then push coupons based on that correlated search directly to her mobile device. And that vision is precisely what is driving the current data bubble, in which online behavioral advertising is overvalued, data brokers calculate the speculative futures of data (hedging bets on the unknown uses to which it will be put), and new computational systems are designed to manage both these speculations and the data sets themselves.

We are thus in the midst of what is exuberantly called a "Data Renaissance," in which new marketing worlds await exploration and raw material—raw data—awaits extrapolation, circulation, and speculation. Data has been figured as a "gold mine" and as "the new oil of the Internet and the new currency of the digital world," the engine driving our latest speculative bubble.⁸ (Around the time of the worldwide financial crash, venture capital began pouring into online tracking.^{γ}) Data speculation means amassing data so as to produce patterns, as opposed to having an idea for which one needs to collect supporting data. Raw data is the material for informational patterns still to come, its value unknown or uncertain until it is converted into the currency of information. And a robust data exchange, with so-termed data handlers and data brokers, has emerged to perform precisely this work of speculation. An illustrative example is BlueKai, "a marketplace where buyers and sellers trade high-quality targeting data like stocks," more specifically, an auction for the near-instant circulation of user intent data (keyword searches, price searching and product comparison, destination cities from travel sites, activity on loan calculators).¹⁰ If the catalog era depended on a stable indexical link between data and subject, the behavioral data banks of the present need repeatedly to enact that link through database operations that are not incidentally termed "join" and "union." In other words, my data does not need to be stabilized as a composite profile subject to the interpretive work of personality analysis and motivation research; what matters is simply its functionality in a particular context at a particular moment. In 1993 Critical Art Ensemble suggested that we might begin to thwart the thenemergent data systems by contaminating them with corrupt or counterfeit data.¹¹ However, data can no longer lose "privilege once it is found to be invalid or unreliable," as they suggest, not only because its truth is operational—if it works it is good—but also because its future value cannot now be calculated. That is, it awaits the query that

()

would produce its value. Data cannot "spoil" because it is now speculatively, rather than statistically, calculated.¹²

The name for the disciplinary and control practice of monitoring, aggregating, and sorting data is dataveillance, named as such by Roger Clarke, who suggested nearly twenty-five years ago that it was then "technically and economically superior" to the two-way televisual media of George Orwell's fictional universe.¹³ It is such because dataveillance operations do not require a centralized system, provided a set of different databases are networked and provided that they share the same means of establishing individual identification, so that a single unit (an individual or number) can be identified consistently across a range of data sets with a primary key. Dataveillance is not new to information technologies and certainly one could construct a genealogy of biopolitical management that would include paper-based techniques such as the U.S. census. Indeed, in an early commentary on the "electronic panopticon," David Lyon suggests that the difference made by information technologies is one of degree not kind, that they simply "make more efficient, more widespread, and simultaneously less visible many processes that already occur."¹⁴ However, one could argue that there have been qualitative as well as quantitative shifts in dataveillance practices in the last decade, or, more precisely, that an intensification of quantitative differences allows for the articulation of qualitative difference. Dataveillance in the present moment is not simply descriptive (monitoring) but also predictive (conjecture) and prescriptive (enactment). To invoke Gilles Deleuze on the emerging structures of continuous control and assessment, "the key thing is that we're at the beginning of something new."¹⁵

The question then becomes: what are the materially distinct features of the new unified and dynamic dataveillance regime? Large-scale data-aggregating corporations such as Acxiom and ChoicePoint and increasingly sophisticated tracking technologies such as Flash cookies and beacons indicate a shift in scale, while the emergence of data exchanges indicate a shift in the evaluation and "appreciation" of data itself.¹⁶The linking of databases, corporate actors, and institutions—as is made possible by corporate acquisitions of DoubleClick (Google) and ChoicePoint (the parent company of Lexis-Nexis)—radically changes the scope of a query, as would the realization of a vision of data storage "measured in petabytes."¹⁷ Speculation lurks here in the incalculable, the size of data storage exceeding conventional metrics and simply open to an unknowable future. Thus is it necessarily the case that data markets should be speculative, their units of exchange not even stabilized as such, and driven by techniques of "predictive optimization" that attempt to generate future value.¹⁸

Data Subjects

The syncing of browser history with personal and application data has successfully and for the most part uncontroversially been situated under the rubric of "enhanced user experience." Apart from the brief bursts of quasi-theatrical collective outrage—we are shocked to hear Google CEO Eric Schmidt remark that "we don't know enough about you. That is the most important aspect of Google's expansion" or to learn of Facebook's creative interpretations of privacy—there seems to have been a general acquiescence to the notion that the distinctions between private and public and personal and nonpersonal when it comes to data are at best tenuous and that it is practically and economically in our interest to regard them as such.¹⁹ Indeed, even as the Wall Street Journal starkly warns its readers to attend to the question of "What They Know," it continues to speculate on the economic growth potential of data mining. The tone and tenor of comments in user forums ranging from Yahoo Answers to Mozilla Support and Computing.net is remarkably consistent: there are basic steps one can take to delete cookies, but it seems unnecessary to do so because they do not interfere with everyday computer use; in fact, some of them are functionally necessary and the end result is that one encounters advertisements that may be of interest. In order to receive customized rather than generalized services, one of course has to provide information to corporations and institutions so that they might better support our preferences, profiles, and favorites. After all, this line of thinking holds, do we not want a personalized Internet that adapts to our individual tastes, habits, and preferences? That it is even possible to speak in such general terms about conditioned behavior is evinced by the memes that play with Google's predictive text feature: What does it think I want when I type "cow"? What does it think my friends want? What mark of distinction accrues to me if the first result is "cowboy bebop" as opposed to "cow clicker"? Such information is shared, circulated, and entered into the field of communicative exchange. In this respect, dataveillance takes its place among affect-generating mechanisms such as Facebook: voluntarily surrendering personal information becomes the means by which social relations are established and collective entities supported. Does this, however, necessarily mean that resignation and ironic acceptance of the new data economy are the doxa?

Pointed questions about behavioral targeting will reveal a certain discomfort from a representative segment of the population; for example, 66 percent of a survey population of adult Americans indicated that they did not want personalized advertising, a number that grew to 73–86 percent when participants were told exactly how

companies collect data for targeted ad campaigns.²⁰ In spite of this, however, the general claim can still be upheld: if in response to the proposed National Data Center in the mid-1960s there was a significant pushback from Congress, the mass media, legal scholars, and the public, in the present moment Americans on the whole seem not to mind being mined.²¹ It might then at first glance seem to be possible to speak, as does Mark Poster, of our "interpellation" by databases. True interpellation-in his terms "a complicated configuration of unconsciousness, indirection, automation, and absentmindedness"-requires a coercive system, a "superpanopticon," capable of rendering us as both subjects of and subjects to that particular assemblage that David Mitchell, in a fictional context, calls a corpocracy.²² For Kevin Robins and Frank Webster, this is the essence of "cybernetic capitalism," by which they mean the whole of the socioeconomic control system that is in part dependent on the capacity of state and corporate entities to collect and aggregate personal data to the extent that each individual can be easily monitored, managed, and hence controlled.²³ As my epigraphs indicate, Robins and Webster are far from alone in their concern with our dynamic incorporation within a totalizing technological system of data management.²⁴ Greg Elmer also explicates the techniques by which consumer profiles are developed and individuals are "continuously integrated into a larger information economy and technological apparatus."25 But for Elmer and Lyon and others, a crucial aspect of this incorporation is our voluntary participation: the composition of consumer profiles in part results from solicitation-whether in the form of a request for feedback or personal data so as to be granted access to a particular service or program—which means we are interpellated as "self-communicating" actors.²⁶ To be sure, to participate in the project of modernity has arguably always meant that one becomes a calculable subject by voluntarily surrendering data. Note the established meaning of "datum" itself as it is recorded in the Oxford English Dictionary: "a thing given or granted; something known or assumed as fact, and made the basis of reasoning or calculation." In the specific context of a sociotechnological milieu organized according to the operational principles of "cybernetic capitalism," however, our acts of participation or self-communication themselves become data, the entirety of our everyday life practices subject to, and constituted by, perpetual calculation. What was speculative at the time of Don DeLillo's White Noise (1985)—"you are the sum total of your data"—has in the intervening years become actualized, and neither the legal nor the political infrastructure has kept pace with the technology.²⁷

In December 2009, Google announced that search would thereafter be personalized according to fifty-seven signals, among them location, machine and browser information, and prior search history.²⁸ The company soon assured its users that it was "recognizing your browser, not you," but who or what is meant by "you" in this formulation? In one account, the "you" is our "data double." Kevin D. Haggerty and Richard V. Ericson explain:

Surveillance technologies do not monitor people *qua* individuals, but instead operate through processes of disassembling and reassembling. People are broken down into a series of discrete informational flows which are stabilized and captured according to pre-established classificatory criteria. They are then transported to centralized locations to be reassembled and combined in ways that serve institutional agendas. Cumulatively, such information constitutes our "data double," our virtual/informational profiles that circulate in various computers and contexts of practical application.²⁹

Financial, travel, and governmental databases might be coordinated but our "data doubles" are only temporarily aggregated, our user profiles produced as an effect or consequence of search queries rather than preexisting stable entities that are then subject to search. It is at this point then that the interpellation argument falters because the processes of subjectification at the heart of the "panoptic sort" have been transformed. Along the same lines, Matthew Fuller argues that surveillance is no longer about visual apprehension but is instead a "socio-algorithmic process" that captures and calculates "flecks of identity," the data trails of our everyday actions, such as our browsing history, financial transactions, and our movements as they are recorded by GPS coordinates on our mobile devices and RFID tags in passports and identity cards.³⁰ The "flecks" concept emerges in some respect from Gilles Deleuze's outline of the emergence of the "dividual" in the context of the control society; if the individuated self was both product and figure of modernity, "dividuals" are rather fragmented and dispersed data bodies. They are, as Tiziana Terranova explains, "what results from the decomposition of individuals into data clouds subject to automated integration and disintegration."31 Put another way, they are the CDOs (collateralized debt obligations) of the data market, in which bits and pieces of a supposed composite profile, which is itself an operative fiction, are sliced and diced into different tranches, such that a stable referential link to a singular entity becomes lost in a sea of user intent data. The noworthodox market position is that the value of data does not depend on its connection

8/3/2012 9:18:44 AM

to an actual person, until expedience requires that a claim be made for the truth of that data. Our data bodies then are repeatedly enacted as a consequence of search procedures. Data is in this respect performative: the composition of flecks and bits of data into a profile of a terror suspect, the re-grounding of abstract data in the targeting of an actual life, will have the effect of producing that life, that body, as a terror suspect.

Countervailing Engagements

Jack Gladney, the principal character in White Noise, is exposed to an airborne toxin and thereafter subjected to a battery of medical tests. The test results are then aggregated with all of his genetic, civic, and personal information to produce a "massive data-base tally," the source and physical location of which are not identified.³² Gladney considers the conspiratorial implications: "I wondered what he meant when he said he'd tapped into my history. Where was it located exactly? Some state or federal agency, some insurance company or credit firm or medical clearinghouse?" No mere paranoid fantasy, the idea of a single national data center as a matter of public policy was considered during congressional hearings in 1966, with technocratic efficiency weighed against civil liberties, specifically the right to privacy, and a number of representatives expressing concern about the fact that "the computer neither forgives nor forgets" and is "incapable of making allowances for early errors or indiscretions."33 As Paul Ohm has proven with careful detail, this exact vision of a data bank that "neither forgives nor forgets." Is in theory realizable because of reidentification—the reversal of anonymization techniques with such relative ease that anonymization cannot and should not be considered a means of privacy protection.³⁴ Perfect anonymity is impossible, but the nightmare scenario (then and now) imagines a womb-to-tomb "record prison" or "database of ruin," a massive "database in the sky" held by Google or elsewhere that contains the material necessary to reduce the entropic uncertainty about individual identities and thus cause demonstrable and legally recognized harm to everyone recorded within it. Google's incorporation of DoubleClick, one of the largest behavioral targeting companies, as well as its partnership with Verizon, would likely be the closest approximation of this single database fantasy, but there is as yet no one entity legally (and technologically) capable of aggregating the entirety of "our" data, which would include not only all governmental and financial records but also our entire search and purchase history, along with our relationship to the social graph. (The value at present is in the aggregating of just a few of these data components.) It is the more general sense that data storage is permanent

that leads Viktor Mayer-Schönberger to claim that we have been produced as Borgesian figures, Funes who have lost the capacity to forget and thereby lost the capacity to structure a temporal narrative.³⁵ More concretely, the consequence of total storage is that the much-heralded second act of American lives—the mythology of reinvention—cannot be possible if all of the data from the first act is easily accessible.

Data storage of this scale, potentially measured in petabytes, would necessarily require sophisticated algorithmic querying in order to detect informational patterns. For David Gelernter, this type of data management would require "topsight," a topdown perspective achieved through software modeling and the creation of microcosmic "mirror worlds," in which raw data filters in from the bottom and the whole comes into focus through statistical modeling and rule and pattern extraction.³⁶ The promise of topsight, in Gelernter's terms, is a progression from *annales* to *annalistes*, from data collection that would satisfy a "neo-Victorian curatorial" drive to data analysis that calculates prediction scenarios and manages risk.³⁷ What would be the locus of suspicion and paranoid fantasy (Poster calls it "database anxiety") if not such an intricate and operationally efficient system, the aggregating capacity of which easily ups the ante on Thomas Pynchon's paranoid realization that "everything is connected"?³⁸

Happily, sheer impracticality means that data systems can never function as perfectly as our dystopian imaginations might suspect. The errors inherent within a catalog mailing list, one of the more basic datasets, indicates how unstable that data can be: any given population is a massive moving target, all the more so considering the inevitable introduction of false information, and the scale of the sample size—in the TIA topsight scenario, for example, every human entity within the U.S. borders—means that it truly would require the storage of petabytes of data in order to produce accurate calculations. Even if one were to accept the fiction of the universal database managed by a single authority, the fundamental problem of meaningfully, and predictably, parsing that archive remains. Everything might be collected and connected, but that does not necessarily mean that everything can be known. Google may come to possess the sum total of my personal data and all of the history contained within my UID, but it cannot obtain the programmatic perspective necessary to predict exactly what I will buy or what I will read.

Still, as my Firefox add-on, Collusion, reminds me, data collection companies are continually tracking my browsing behavior in spite of my efforts to thwart them, a cogent reminder that targeting is not impractical at the level of the individual. When considered in these terms, it is difficult to dismiss escape, whether in the form of

disappearance or disconnectivity, as merely a counterfantasy.³⁹ Critical Art Ensemble's injunction is to the point: "Avoid using any technology that records data facts unless it is essential."40 Howard Rheingold and Eric Kluitenberg make a comparable case for "selective connectivity": techniques by which we can "choose to extract ourselves from the electronic control grid from time to time and place to place."⁺¹ Similarly, for Mayer-Schönberger, the solution lies in the adoption of a certain care in the management of one's online interactions, practices of selective disclosure, and revelation in order to limit "uncontrollable information flows through individual choice."⁴² If we are able to opt out of a single company's personalized retargeting scheme, that is, should we not also be able to opt out of all advertising databases or indeed out of the whole system of "cybernetic capitalism" itself? But it is arguably the case that exit in the form of forgetting or genuine anonymity is no longer possible, that disappearance itself has disappeared. Confronted with this argument we might instead imagine a systems overload, "an information blizzard—a whiteout," because silence can be attained with an increased pitch of white noise.⁴³ "Anonymity systems function best in a crowd" and therefore overflowing the system, feeding it false information, generating more "flecks of identity" than it can handle, might be the closest approximation of disappearance it is possible to achieve.⁴⁴ A creative example of precisely this is Daniel Howe and Helen Nissenbaum's TrackMeNot, a browser extension that works to block the capacity of third parties to identify users based on their search history by periodically creating bursts of search activity and thus hiding real searches within a batch of ghost queries. As the creators explain: "To level the playing field, we have sought to create a mechanism that places some degree of control back in the hands of users and, at every point in the design where this has been feasible, we have sought to do so."45 Counterpropositions such as these, however, shift the burden of governance from institutions to the mythic entity of the individual rational actor and either argue for or presume a certain technological literacy from the outset.⁴⁶ They also imply that data is somehow neutral and that it is only the uses of data that is either repressive or emancipatory.

The critical minefield one must negotiate here is structured by two tried-and-true narratives: one outlining systems of control and the other positioning us as well-informed citizens who can manage (indeed "give") our data and perhaps even turn dataveillance techniques to our own advantage. The version of this binary particular to the Internet pits monopolistic corporations seeking jurisdiction over information architecture and communication flows against those fighting to maintain open, distributed P2P networks (Google is the complicating exception in that it is a single entity whose

power derives from the management, support, and ownership of those very distributed networks). If considered in narrowly exclusive terms, each narrative risks a certain blindness: either an overinvestment in the valorization of the agency of the user who hacks the system or an overinvestment in the articulation of the protocols of a given system as inescapably binding, such that it would require naively idealistic faith if not false consciousness to believe in the efficacy and value of resistant and participatory practices. But it remains the case that constellations of control are imbricated with constellations of expressive resistance, whether in the form of tactical intervention, asymmetric infowar, or civic engagement. For every system of disciplinary power, as Anthony Giddens puts it, there is a "countervailing" response from those in precarious, subordinate, or marginal positions, which is to say that dataveillance and countervailance must be seen as inextricably connected.⁴⁷ The practices that might be situated under the rubric of countervailance do not endeavor to realize an abstracted philosophy of resistance and human rights. They are often cognizant of such rights, particularly when a governmental program like Poindexter's TIA is articulated within the field of tactical activity as a critical object. But their actions are more often about action itself in relation to a regime that would limit us to efforts to stay on the right side of the data that defines us. Moreover, the expressive aspects of countervailance as I will outline them here serve as an important counter to the technocratic consumer rights initiatives that frame the debate in terms of property-those "MyData" initiatives that seek only to transfer ownership of data to the individual and to develop personal data banks for everyday functionality and monetization.⁴⁸

There are a number of practices that have the potential for disruptive innovation vis-à-vis the new regime of dataveillance. For example, Gary Marx outlines a range of behavioral techniques and legal, economic, and technological exploits ranging from refusal to masking that work toward "neutralizing and resisting the new surveillance" system; neutralization, as he puts it, is a "dynamic adversarial social dance involving strategic moves and counter-moves and should be studied as a conflict interaction process."⁴⁹ With respect to consumer (re)targeting and behavioral profiling, a common counter-move is the design and programming of anonymizers, encrypters, distributed networks, and ad and cookie blockers. Though many such enterprising programmers may work for large IT corporations, their software can usually be tagged as independent, alternative, open, and almost always free. Just as Internet data mining is dependent on software design, then, so, too, is the blocking or thwarting of that mining. So, to block beacons and zombie cookies and maintain the smallest measure of privacy while reading

8/3/2012 9:18:44 AM

articles in the *Guardian* online, one can choose from a suite of effective Firefox add-ons including TACO and Beef TACO (targeted advertising cookie opt-out); BetterPrivacy; Ghostery; CookieSafe; and CookieCuller. As Panopticlick, the Electronic Frontier Foundation's browser-fingerprinting algorithm, reveals, however, privacy tools such as spoofers and plugins paradoxically make the browser more distinct and thus facilitate device fingerprinting.⁵⁰ Panopticlick further reminds us of the difficulty of demarcating an absolute difference between the means of tracking and the means of circumventing that tracking; another case in point would be browsers in which the facility for private browsing is built into the browser itself.

To understand the significance of software design both to mine and to obstruct, one has only to consider the role that computer models have played in what Andrew Leyshon and Nigel Thrift describe as "the capitalization of almost everything," which is to say in the creation of the explosive development of financial capitalism that led up to the recent global financial crash. In short, "new forms of expertise, fuelled by computing power and software" have been necessarily constitutive.⁵¹ For example, the consolidation and centralization particular to "Shared Services" would not have been possible without the development of a single operating system to aggregate a range of different activities and income streams into a single entity. "As in the case of ground rent, what made the mining of these new seams of financial value [subprime lending] apparently possible is the development of computer software that enables individuals to be assessed, sorted and aggregated along dimensions of risk and reward."52 Software design did not simply enable the creation of new financial instruments; software design was the essential condition of possibility for these new financial instruments. So, too, the scale and complexity of the data structures at issue—"petabytes"—is such that they cannot be processed by human intelligence alone but rather require machine intelligence in the form of database management systems and algorithms that structure data collection.

Combating or otherwise responding to a control system dependent on computing power requires the design of a counter-system, a rather modest example of which is Diaspora, an open-source, privacy-aware, distributed do-it-yourself social network that eliminates the hub of a social media conglomerate in favor of a peer-to-peer network in which each individual is a node.⁵³ Without a hub or central server, data encrypted with GNU Privacy Guard is sent directly to one's friends rather than stored and hence mined. True peer-to-peer communication—that is, that which is not routed through a central hub—would need to move to a network such as Diaspora because the controlled application programming interface (API) of social networks such as Facebook means that

hacking a hub-based network in order to convert it to peer-to-peer is difficult if not impossible. Regardless, we could point to numerous examples of Facebook users harnessing the peer-to-peer over central hub to mobilize street-based protests, in essence modifying a digitally centralized network so that it functions as peer-to-peer.

Another common countervailing response has been the appropriation of the technological tools of surveillance-whether that be "reciprocal transparency" (watching the watchers) or lateral surveillance, the myriad ways in which people keep track of each other with social networking platforms, cameras, and GPS-enabled mobile devices.⁵⁴ Indeed, in the context of social media, lateral surveillance has been considered as a sharing practice involving mutuality and reciprocity rather than a one-way flow of information.⁵⁵ So, too, self-directed profiling ("my preferences") means articulating one's own value as a consumer, traveler, citizen, and friend. While dataveillance functions as an instrument of biopolitical control, in other words, it also enables civic participation, at least insofar as one regards as significant the effects of private citizens performing both their own background checks with Google and Facebook and their own market research through user ratings and sites such as Yelp. "Folksonomies," usercreated systems for establishing value (via tagging, bookmarking, and rating) similarly function as a means of making community. From Amazon to Digg, there is a vast network to which we can turn to assess our value as producers (of comments, reviews, commodities) and consumers (as trusted users and buyers), one whose seemingly inconsequential rewards (stars, levels) mask a deep sense of community. In this respect, making data public is also making a commons. Apart from functioning as a rival form of expertise, then, one effect of these countervailing tools and techniques has been to re-embed dataveillance within social relations. Perhaps the best example of this is Eyebrowse, a protosocial network based on the self-reporting of one's browsing activities (figure 7.1). A Firefox plugin, Eyebrowse visualizes a user's web browsing history along with that of her friends, thus making visible the data available to Google and any number of third parties, now and in the future.³⁶

Mimetically reproducing data collection practices increases technological literacy with respect to both individual everyday practice and systemic logics. Exploiting vulnerabilities makes those vulnerabilities known. Evercookie is perfectly illustrative. The virtuosic work of an elite hacker, evercookie is as it sounds, a tracking device that cannot be destroyed. Designed as a "litmus test," with the tag line "never forget," evercookie provides incontrovertible proof of our relative inability to control the storage of cookies on our computers, particularly in the scripting environment of HTML 5.⁵⁷ A more

8/3/2012 9:18:44 AM

()

êyêbrowse

۲

track, visualize and share your web trails

Eyebrowse is an add-on for <u>firefox</u> that records your web browsing activity to your private computer so you can check out how your activity changes over time. Selectively share your activities to find out what's hot and who's reading what.



Figure 7.1 Eyebrowse, created by Brennan Moore, Max Van Kleek, and David Karger (MIT CSAIL).

ordinary example is the Firefox extension Firesheep, which allows users to capture the unencrypted login cookies of others on the shared Wi-Fi network, thereby substantiating the need for HTTPS. The hope is that participatory and educative tracking tools such as these produce a more-informed public and blur the lines between a data class that does not understand at a basic level how cookies function and a class of power users savvy enough to exploit the resources at their disposal in the interests of constituting their own data bodies. What becomes apparent after several hours of hands-on work tinkering in search of the perfect combination of antitracking tools, however, is that expert knowledge quickly becomes the aspirational goal, with legal and technological complaints about data mining mollified by the temporary satisfaction of having joined the elite data class. Nonetheless, an embodied experience of dataveillance tools and techniques alerts the public to its role as a stakeholder for, Alberto Melucci notes, "as mere consumers of information, people are excluded from the discussion on the logic that organizes this flow of information; they are there to only receive it and have no access to the power that shapes reality through the controlled ebb and flow of information."58 A tool such as Eyebrowse certainly gives its users access to data collection

processes, though it might well introduce the question of the extent to which we are being asked to immerse ourselves in the dataveillance regime to the point of identification in order to achieve any sort of agential position. Because inhabitation prompts recognition, however, a fully immersive, participatory, and identificatory practice can still function as a means of using a control apparatus against itself.

Mirror Worlds

Artists who appropriate dataveillance techniques and tools as a medium for creative production inform, enlighten, and help us to imagine otherwise by refusing the fantasy of exodus, a withdrawal from a given political, economic, or cultural system predicated on the notion that there is a neutral external vantage point from which one can begin the work of critical assessment.⁵⁹ In a very general sense we might term such work immanent critique: art-activism operating within a given structure and inhabiting a particular perspectival frame, whether that be bioartists' hands-on work in the laboratory or hacktivist interventions within networked systems. The paradigmatic instance of an art practice that inhabits a particular perspectival frame would be that of the Yes Men, whose counterfeit performances in the name of entities such as the WTO, Halliburton, and Dow Chemical continue to be mistaken for the real. In work such as this, critique is situated in the act of mimesis, which is not a refusal of "corpocracy" but a reflection in a double sense: mirroring and replication, on the one hand, and critical contemplation on the other. A reiterative aesthetic serves to engage a public with a reflective understanding of the operations of power and control. Its creative, productive, and playful aspects open rather than foreclose lines of inquiry; in its eschewing of a singular and reductive negative judgment, it maintains a purchase on continuous critical assessment. A reiterative aesthetic can be radically transformational precisely because it exists in dynamic interplay with its object; it neither claims a stable outside nor fixes upon a synchronic slide of a system that is the inevitable byproduct of topsight.

The work of the Preemptive Media collective—whose practice includes instructional workshops and the re-engineering of mobile technologies—is particularly apposite for a discussion of dataveillance and tactical countervailance. Preemptive Media's object is to exploit consumer electronics for a larger purpose, to foster not only technological literacy, but also critical consciousness and a kind of low-tech amateurism. In one representative series of performances, called *SWIPE* (2002–2005), the collective installed a functioning bar in galleries and exhibition spaces and opened it up for enjoyment and

8/3/2012 9:18:44 AM

play.⁶⁰ Patrons ordering drinks had their drivers licenses scanned and were given individual receipts detailing the data culled both from the 2D barcode and online search. Computer stations in the bars displayed a web-based toolkit with a data calculator to allow participants to determine the market value of their individual data; they also displayed the decoding application used in the installations and a thorough guide to the process of requesting one's data files from the large data warehouses: ChoicePoint, Acxiom, LocatePLUS, and Experian. The purpose was to encourage consumer awareness of Automated Identification Data Capture technologies (AIDC); to give participants the experience of visualizing their own data; and to facilitate a critical conversation about data mining, transparency, and privacy. Swiping suggests purchasing, as if one uses currency to establish or prove currency, reminding us of the extent to which the value, significance, and indeed existence of the individual body are calculated, even proved, by complex systems of accounting-the precise operation of which remains obscure. But SWIPE interrupted the one-way flow of information from evidentiary subject to surveillance mechanism, enacting in the process lateral relations among the participants. As the bar setting indicates, SWIPE functioned within a social space, its relational aesthetic true to Nicolas Bourriaud's vision of an artistic praxis that struggles against the reifying and commodifying of social relations by creating a space for "alternative forms of sociability."61 Even as it introduced a certain defamiliarizing shock in individual participants, then, it was unambiguous about the situation of those participants within a broader political and socioeconomic matrix. As the artists noted: "Our hope is to encourage thinking beyond the individual self ('I do not care if a bar database has my name and address and time of visit . . .') toward understanding databases as a discursive, organizational practice and an essential technique of power in today's social field."62

Osman Khan's installation *Net Worth* (2004) was similarly dependent on the gallery visitor's swipe, in this case of a credit or ATM card, in order to mine the identificatory information necessary to perform a Google search to determine search rank and thus, "net worth"⁶³ Drawing on the familiar practice of egosurfing, the tracing of one's own virtual-physical presence and presumed importance online, this installation articulated a shift from the moment of the televisual record—you don't exist unless the entire world sees your image—to the moment of the database record—you don't exist unless you appear on Google. So, too, *Net Worth* invokes the discourse on reputation and trusted users in its equating of the assessment of net presence with the assessment of the value of the individual. More recently, David Kemp asked 100 people to show him

the identification, banking, and loyalty cards in their wallet—"anything that connects to a database"⁶⁴—and then for his installation, *Data Collection* (2010), he used each data set to compose an individual "canvas" with photographic representations of the cards on which all of the personal information is visible, with some cards blacked out on request of participant. A small sampling of dataveillance art, these projects are both tactile and rhetorical, dependent on the gift of data in order to open a space for the critical contemplation of that data. They work with—both exploit and capitalize upon participants' willingness to share data for no immediately tangible or concrete reward, that is, for no apparent return on their affective and participatory investment. What is illuminated by each is the logic of social media and relational aesthetics, which is to give by sharing.

A skeptical viewer might ask whether such data works are in fact supportive of, and thus insufficiently attentive to, their own corporate and governmental information architecture. But this is a variant of the old worry about artists not having sufficient critical distance from the capitalist, technological, scientific, and ideological systems within which they are working. In other words, to suggest that using data-mining techniques to produce art necessarily entails adopting the very logic and optics of the dataveillance society is to rehash the old problem of disinterest. The common assumption is that distance is necessary for critical reflection and that proximity necessarily produces corruption. But the lesson I think we need to learn from tactical media practitioners more broadly is that critique and critical reflection are at their most powerful when they do not adopt a spectatorial position on the (putatively neutral) outside, when they do not merely sketch a surface, but rather penetrate the core of the system itself, intensifying identification so as to produce structural change.⁶⁵ Such a practice—such a mode and positioning—goes well beyond Michel de Certeau's notion of "undermining a system from within"; these are not employees wasting time and using the resources of the workplace to turn it against itself.⁶⁶ Rather, these art-activists are creating "mirror worlds," replicating the scene of data mining—swiping an identification card—to enable an embedded and embodied perspective of the control network through and within which dataveillance operates, but without the fantasy of exteriority. Instead, the force of the immanent critique envisioned here derives from a near-total inhabitation of the frame, compelling a jarring recognition from the viewer/user and leading to a temporal interval in which she must formulate a response, whether that be rejection or acquiescence. Interventionist art projects such as these work directly against the forces of interpellation with a counterimage of a dataveillance regime that makes that regime

8/3/2012 9:18:45 AM

perceptible—and if it is perceptible then it becomes possible to work concretely toward political transformation.

The role played by the designers of countervailing tactics, tools, and techniques is akin to that played by the "Keymaker" in *The Matrix Reloaded*: they offer access to a back door, a shortcut key or authenticating token that holds out the promise of allowing us to circumvent the programmed structure of the dataveillance regime.⁶⁷ The film is reflexively archetypal. The Oracle instructs Neo, the One, to find the Keymaker, who is being held captive by a master program because of his knowledge of the rules of the system and his ability to open a door leading to The Source. His pre-scripted function is to sacrifice himself to The Resistance project. When Neo opens the door to his prison to find him in the act of making the single key, he announces his function: "I'm the Keymaker. I've been waiting for you." He tells the skeptical ship commanders that he knows of the door and the building level "where no elevator can go, where no stair can reach" because he "must know" and it is "his purpose," the "reason" he is there. And when he is killed by the agents after opening the door to the antechamber, he tells Neo and Morpheus simply that "it was meant to be." In other words, he is programmed only to exploit the weakness in the system, after which he becomes expendable. Read representationally, the Keymaker program is an integral component of the matrix: control systems must necessarily have moles who can reveal the means of puncturing the system so as to satisfy the demand for breaking through (or leveling up)—a demand that is at once narratological and psycho-social. These acts have precise actors ("only The One can open the door"), precise spatiotemporal coordinates ("only during the window can that door be opened"), precise organizational logics ("All must be done as one. If one fails, all fail"), and they can be performed exactly once. Once the door is opened or the threshold crossed, the act cannot be repeated. The flip side of the fantasy of total information awareness, then, is the fantasy of breach.

But the Keymaker does not need narrative structure to legitimate his energies; indeed he dies even before the plot of which he speaks is realized. His role does not exactly duplicate that of countervailing actors—I am not after all advocating sacrifice but it is emblematic. On the one hand his knowledge is scripted ("I know because I must know") and his circumvention of the system thus simply an exercise in selfregulation. The extra-institutional spaces, here the hallway that is not legible within the matrix, are themselves built into the system and subject to management. On the other hand, however, the wily Keymaker does elude the agents and open the door, which is to say that the act is neither a protocol nor sabotage but both, and self-reflexively so. So, too, evercookie, the indestructible cookie, is neither purely a tracking technology nor a hack designed to show vulnerabilities and *SWIPE* is neither actual data collection nor a performance of the same but both/and. In other words, dataveillance and countervailance coexist not in dialectical struggle but rather are so fundamentally entangled that the line separating the one from the other is unstable. Positioned as we are within the dataveillance regime, we cannot but employ the tactics of immanent critique, which depends not on an overstatement or overarticulation of totalizing control systems nor on a hyperbolized romance of the exploitation of these systems, but rather depends simply on ordinary action itself.

Acknowledgments

Many thanks to Russell Samolsky, Lisa Gitelman, Rahul Mukherjee, Juliette Cherbuliez, and Francisco J. Ricardo for careful reading and helpful suggestions. Earlier versions of this paper were presented at the American Political Science Association annual conference in Washington D.C.; "DIY Citizenship: Critical Making and Social Media," University of Toronto; and Department of Communication Studies, Concordia University. Thanks to Renee Marlin-Bennett, Megan Boler, Matthew Ratto, and Charles Acland for the invitations to present this work and to the audiences for significant feedback.

Notes

1. "Search leakage" is the disclosure of search terms to visited sites; that is, a record of the path followed to land on a particular page. Search engines that allow one to surf anonymously, most of which neither record IP addresses nor use identifying cookies, include Scroogle, Ixquick, DuckDuckGo, and Yauba. Another way to prevent search leakage is to use network routing software like Tor, an "infomediary" that encrypts traffic between the individual user and the Tor network. More simply, encrypted search (HTTPS, or HTTP secure) does not send search terms. The Electronic Privacy Information Center (EPIC) maintains an extensive list of privacy tools for voice, email, instant messaging, and browsing, as does the Center for Democracy and Technology. See http://epic.org/privacy/tools.html and https://www.cdt.org/privacy/guide/basic/tips.php (accessed February 7, 2011).

2. A 2009 article in *Wired*, admittedly usually a bit delayed both with its techno-boosterism and techno-paranoia, suggests that LSOs have for the most part escaped general notice, a point made in a number of related articles then and since. See Ryan Singel, "You Deleted Your Cookies? Think Again," *Wired* (August 10, 2009), http://www.wired.com/epicenter/2009/08/you -deleted-your-cookies-think-again (accessed August 10, 2009). By the time this chapter makes it into print, it, too, will likely seem a bit belated, particularly as HTML 5 comes into widespread

<u>R</u>

use, but it can be read as a snapshot account of dataveillance practices and the tactics, techniques, and technologies deployed to negotiate them in the era of big data, a battle that will almost certainly persist for the foreseeable future.

3. Adobe Statement for the Privacy Privacy Roundtables Project filed with the Federal Trade Commission (January 27, 2010), http://www.ftc.gov/os/comments/privacyroundtable/544506-00085.pdf (accessed July 25, 2010). Clearspring Technologies, one of the larger content-sharing companies, and the developer of the AddThis platform, discloses its use of Flash cookies in its privacy policy for AddThis, but not on the privacy policy for the company itself. See http://www.addthis.com/privacy[REMOVED HYPERLINK FIELD] (accessed November 14, 2010).

4. At the time of this writing, the "What They Know" section of WSJ.com continues to be regularly updated. http://online.wsj.com/public/page/what-they-know-digital-privacy.html (accessed February 7, 2011).

5. The concepts of "personal" and "nonpersonal" are, as one would expect, somewhat mutable in the context of dataveillance. The single cookie assigned to each machine is not automatically attached to an individual identity so, while sexual preference might in certain legal statutes be defined as "personal," in the context of information security it would be considered nonpersonal. Personally identifiable information (PII), on the other hand, includes social security numbers, genetic information, biometric data, date of birth, and in some cases vehicle registration numbers, bank numbers, and IP addresses, although the increasingly widespread use of proxies makes the last more complicated. Much of the data-privacy legislation to date restricts the use of PII and presumes the safety of anonymization.

6. The Adobe AudienceManager platform, which is based on Demdex, invites companies to create a data bank based on both their own ad campaigns and data acquired from third parties. http://www.demdex.com (accessed February 10, 2011).

7. John Battelle, *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture* (New York: Penguin, 2005), 6. The structural logic behind online behavioral advertising would be the "panoptic sort," Oscar Gandy's descriptive formulation for the system that "operates to increase the precision with which individuals are classified according to their perceived value in the marketplace and their susceptibility to particular appeals." Oscar H. Gandy, *The Panoptic Sort: A Political Economy of Personal Information* (Boulder, CO: Westview Press, 1993), 2.

8. Julia Angwin, "The Web's New Gold Mine: Your Secrets," *Wall Street Journal* (July 30, 2010). Meglena Kuneva, cited in Marc Davis, keynote presentation, Privacy, Identity, Innovation annual conference (Seattle, 2010), http://vimeo.com/14401407 (accessed November 12, 2010).

9. Scott Thurm, "Online Trackers Rake In Funding," *Wall Street Journal* (February 25, 2011), http://online.wsj.com/article/SB10001424052748704657704576150191661959856.html (accessed November 12, 2010).

10. See http://www.bluekai.com (accessed November 12, 2010).

11. Critical Art Ensemble, The Electronic Disturbance (New York: Autonomedia, 1993), 63.

12. Ibid., 140.

13. Roger Clarke, "Information Technology and Dataveillance," *Communications of the ACM* 31, no. 5 (May 1988): 499; http://www.rogerclarke.com/DV/CACM88.html (accessed November 12, 2010). In this chapter I focus specifically on dataveillance in the sense of data mining (capture and aggregation), as opposed to the whole suite of techniques and technologies of a contemporary electronic surveillance regime, ranging from CCTV to biometrics, though they are by no means unrelated.

14. David Lyon, The Electronic Eye: The Rise of Surveillance Society (Minneapolis: University of Minnesota Press, 1994), 40.

15. Gilles Deleuze, "Postscript on Control Societies," *Negotiations, 1972–1990*, trans. Martin Joughin (New York: Columbia University Press, 1995), 182.

16. Beacons such as web bugs and pixels track user keyboard and mouse activity on a given webpage.

17. Cited in Elliott Borin, "Feds Open 'Total' Tech Spy System," *Wired* (August 7, 2002); see http://www.wired.com/politics/law/news/2002/08/54342 (accessed August 25, 2010).

18. One company, [x+1], has named its product the Predictive Optimization Engine (POE TM). See http://www.xplusone.com/glossary (accessed November 10, 2010).

19. Quoted in Ira Rubinstein, Ronald D. Lee, and Paul M. Schwartz, "Data Mining and Internet Profiling: Emerging Regulatory and Technological Approaches," *University of Chicago Law Review* 75 (2008): 273.

20. We can say, then, that the privacy crisis produced by the new practices of data collection is to a certain extent hidden in plain sight and recognizable only in moments of elucidation. Joseph Turow, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy, "Americans Reject Tailored Advertising and Three Activities That Enable It" (September 29, 2009). Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1478214 (accessed November 10, 2010).

21. Arthur Miller provides a full account of the proposal and its reception in his seminal text, *The Assault on Privacy: Computers, Data Banks, and Dossiers*, which was itself positioned as a response to the idea. As he notes, the public debate managed to avoid "the fundamental policy issue of how to curtail the government's increasing penchant for information collection" and a defeat of the proposal for the national data network simply meant that each governmental agency developed its own: *The Assault on Privacy* (Ann Arbor: The University of Michigan Press, 1971), 59. In a quite general sense, public reaction needs to be situated in the particular historical,

()

sociocultural, and juridical contexts of a nation-state. Consider, by contrast, the German government's successful campaign against Google's Street View feature. For an early report on the issue, see Kevin O'Brien, "Google Data Admission Angers European Officials," *NewYork Times* (May 15, 2010), http://www.nytimes.com/2010/05/16/technology/16google.html (accessed May 15, 2010).

22. Mark Poster, "Databases as Discourse; or, Electronic Interpellations," *Computers, Surveillance, and Privacy*, ed. David Lyon and Elia Zureik (Minneapolis: University of Minnesota Press, 1996), 187. The Sonmi chapters of Mitchell's *Cloud Atlas* are set in the dystopian corpocratic state called Nea So Copros.

23. Kevin Robins and Frank Webster, "Cybernetic Capitalism: Information, Technology, Everyday Life," *The Political Economy of Information*, ed. Vincent Mosco and Janet Wasko (Madison: University of Wisconsin Press, 1988), 44–75.

24. For a thorough account of pattern-based searches by government and corporations and the techniques one can use to mask online activity; a detailed overview of the public outcry over TIA, its subsequent de-funding, and the continuation of the same data-mining exercises under the classified intelligence budget; and, finally, a detailed legal review that makes the case for transparency and new identity technologies with privacy protections, see Rubinstein, Lee, and Schwartz, "Data Mining and Internet Profiling." It is important to note, however, that these debates are premised on a notion of privacy with a particular history and cultural specificity.

25. Greg Elmer, Profiling Machines: Mapping the Personal Information Economy (Cambridge, MA: MIT Press, 2004), 17.

26. Lyon, The Electronic Eye, 52.

27. There are substantive juridico-political questions that need to be addressed as the legal infrastructure develops: What is the legal status of our financial records, unique ID codes, and biometric data? How or to what extent will individual data be monetized? Can individual browsing be considered labor? If so, would not the unique ID code that records sites visited be considered a product of that labor and thus private property? Does the person from whom data originated have claims over it once it enters into circulation on the "data exchange"? Will data follow the model of genetic materials, with data becoming the intellectual property of a data broker who had altered it in some fashion? Proposed policy solutions thus far include improved securitization, transparency and informed consent, expiration dates and storage limits, and the regulation of data centers.

28. On the era of personalization, see Eli Pariser, *The Filter Bubble: What the Internet is Hiding from You* (New York: Penguin Press, 2011).

29. Kevin D. Haggerty and Richard V. Ericson, *The New Politics of Surveillance and Visibility* (Toronto: University of Toronto Press, 2006), 4.

()

30. Matthew Fuller, *Media Ecologies: Materialist Energies in Art and Technoculture* (Cambridge, MA: MIT Press, 2005), 149.

31. Tiziana Terranova, Network Culture: Politics for the Information Age (London: Pluto Press, 2004), 34.

32. Don DeLillo, White Noise (New York: Viking, 1985), 140.

33. Hearings on the Computer and Invasion of Privacy before a Subcommittee of the House Committee on Government Operations, 89th Congress, 2nd Session (Washington: U.S. Government Printing Office, 1966) 3, 12.

34. Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization," University of Colorado Law Legal Studies Research Paper No. 09-12 (August 13, 2009). Available at SSRN: http://ssrn.com/abstract=1450006 (accessed August 25, 2010).

35. The reference here is to the Jorge Luis Borges story, "Funes the Memorious." Viktor Mayer-Schönberger makes a case for putting an expiration date on information, which would mean customizing each data element so that it remains accessible for a limited period of time—that is, for flexible implementation rather than a general legal code. See *Delete: The Virtue of Forgetting in the Digital Age* (Princeton, NJ: Princeton University Press, 2009).

36. David Gelernter, *Mirror Worlds: Or, the Day Software Puts the Universe in a Shoebox . . . How It Will Happen and What It Will Mean* (New York: Oxford University Press, 1992). John Poindexter's plans for the Total Information Awareness Program (TIA) drew on Gelernter's paradigm, endeavoring to use the principle of topsight to establish a terror network that could ostensibly be seen and disciplined, though not eliminated because of its regenerative ends.

37. Ibid., 112.

()

38. Thomas Pynchon, Gravity's Rainbow (New York: Penguin Books, [1973] 1995), 703.

39. See Irving Goh, "Prolegomenon to a Right to Disappear," *Cultural Politics* 2, no. 1 (March 2006): 97–114.

40. Critical Art Ensemble, *Electronic Disturbance*, 135.

41. Howard Rheingold and Eric Kluitenberg, "Mindful Disconnection: Counterpowering the Panopticon from the Inside," *OPEN 11 Hybrid Space* (Amsterdam: NAi Publishers, 2007), 32.

42. Mayer-Schönberger, Delete, 129.

43. Critical Art Ensemble, *Electronic Disturbance*, 132.

44. Rubinstein, Lee, and Schwartz, "Data Mining and Internet Profiling," 277. It is not uncommon to hear this argument made with respect to social media; in other words, if everyone's intimate details are available, we are essentially hidden in plain sight.

 \bigcirc

45. TrackMeNot FAQ, http://cs.nyu.edu/trackmenot/faq.html (accessed August 25, 2010).

()

46. The detailed counsel about risk management online offered by the Electronic Frontier Foundation's Surveillance-Self Defense Project is paradigmatic. See https://ssd.eff.org (accessed October 14, 2011).

47. Anthony Giddens, *The Nation-State and Violence* (Berkeley: University of California Press, 1985), 186.

48. See Richard H. Thaler, "Show Us the Data. (It's Ours, After All)," *NewYork Times* (April 23, 2011), http://www.nytimes.com/2011/04/24/business/24view.html (accessed April 23, 2011) and "Better Choices, Better Deals," U.K. Cabinet Office (April 13, 2011), http://www.cabinetoffice.gov.uk/resource-library/better-choices-better-deals (accessed July 14, 2011).

49. Gary Marx, "A Tack in the Shoe: Neutralizing and Resisting the New Surveillance," *Journal of Social Issues* 59, no. 2 (May 2003), 369–390.

50. The odds of someone in my time zone using the same browser, operating system, font set, privacy tools, and precise microversions of plugins (Java 1.6.0_17) are remarkably low. Cookies leave crumbs, however dispersed and persistent, while fingerprinting does not, which means that browser tagging essentially goes undetected. Though fingerprints can be associated with search terms, the economic and political utility of fingerprinting, apart from authentication, is not yet entirely clear, which again suggests that what is at stake are potential uses and abuses.

51. Andrew Leyshon and Nigel Thrift, "The Capitalization of Almost Everything: The Future of Finance and Capitalism," *Theory, Culture & Society* 24, no. 7–8 (2007): 101.

52. Ibid., 108.

(�)

53. Developed by four Columbia University undergraduates in response to a lecture on Internet privacy, Diaspora is as of this writing still in alpha version. See http://www.joindiaspora.com and Jim Dwyer, "Four Nerds and a Cry to Arms Against Facebook," *New York Times* (May 11, 2010), http://www.nytimes.com/2010/05/12/nyregion/12about.html (accessed May 11 2010).

54. See David Brin, *The Transparent Society:Will Technology Force Us to Choose Between Privacy and Freedom?* (New York: Basic Books, 1999); and Mark Andrejevic, "The Work of Watching One Another: Lateral Surveillance, Risk, and Governance," *Surveillance & Society* 2, no. 4 (2005), 479–497.

55. Anders Albrechtslund, "Online Social Networking as Participatory Surveillance," *First Monday* 13, no. 3 (March 3, 2008), http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2142/1949#6 (accessed September 22, 2010).

56. See Marx Van Kleek, Christina Xu, Brennan Moore, and David R. Karger, "Eyebrowse: Real-Time Web Activity Sharing and Visualization," *CHI 2010* (April 10–15, 2010). ACM 978–1-60558–930–5/10/04.

57. Tanzina Vega, "New Web Code Draws Concern Over Privacy Risks," *New York Times* (October 10, 2010), https://www.nytimes.com/2010/10/11/business/media/11privacy .html?hp (accessed October 10, 2010).

()

58. Alberto Melucci, *Challenging Codes: Collective Action in the Information Age* (Cambridge: Cambridge University Press, 1996), 180.

59. A full account of surveillance art is necessarily outside the scope of this chapter; on this topic see Thomas Y. Levin, Ursula Frohne, and Peter Weibel's nearly comprehensive *[Ctrl] Space: Rhetorics of Surveillance from Bentham to Big Brother* (Karlsruhe, Germany: ZKM, 2001 (Cambridge, MA: MIT Press). The works I have selected for discussion more narrowly engage the collection of consumer data.

60. *SWIPE*, http://www.preemptivemedia.net/swipe/bar/index.html (accessed February 28, 2012).

61. Nicolas Bourriaud, Relational Aesthetics (Paris: Les Presses du réel, 2002), 44.

62. *SWIPE*, http://web.archive.org/web/20060117165314/http://www.we-swipe.us/plain .html#about (accessed February 28, 2012).

63. Osman Khan, *Net Worth*, http://www.todayandtomorrow.net/2005/09/15/net-worth (accessed September 1, 2010).

64. Quoted in Jan Allen, "Sorting Daemons," Sorting Daemons: Art, Surveillance Regimes and Social Control (Kingston, Canada: Agnes Etherington Art Centre, 2010), 22.

65. I make this case in Tactical Media (Minneapolis: University of Minnesota Press, 2009).

66. Michel de Certeau, *The Practice of Everyday Life* (Berkeley: University of California Press, 1984), 179.

67. *The Matrix Reloaded*, directed by Andy and Lana Wachowski (2003; Burbank, CA: Warner Home Video, 2003), DVD.

()