limn

NUMBER TWO | **CLOUDS AND CROWDS**

Clouds and
↓ Crowds ↑

LIMN NUMBER TWO
CONTENTS | **CLOUDS AND CROWDS**

Edited by Christopher M. Kelty, Lilly Irani, and Nick Seaver

Preface

THIS ISSUE OF LIMN AIMS to raise the level of discussion about new social media, crowdsourcing, cloud computing, big data, and Internet revolutions. Too often, writing about these things follows well-worn paths of argument—paths that become increasingly worn with every rehearsal. The pieces herein seek to interrupt that path, to cross it at odd angles, to find another way through the complex thicket of technology and society.

Take for example the phenomenon known as "Big Data" and the miraculous new forms of problem-solving, knowledge creation and economic productivity it promises. The buzzwords of the brave new world of big data include "cloud computing," "algorithms," "filters," "virtualization," and "scalable infrastructures." Terrabytes and exabytes and petabytes of data

are produced by Facebook and Walmart who analyze them with complicated "machine learning algorithms" and "natural language processing." Breathless claims ("data is in the driver's seat" claims a recent *New York Times* article) and hyperventilatory rhetoric ("The end of theory" claimed *Wired's* Chris Anderson) accompany these developments. The only alternative apparently, is to anxiously and darkly depict world without privacy.[1] The cloud of claims about cloud computing and big data settle into recognizable, if no less nebulous fog banks of enthusiasm or anxiety.

Or consider the Arab Spring of 2011, and the anniversary of the revolution in Egypt this year. The question has repeatedly been posed as to whether the Internet, specifically social media platforms like Facebook and Twitter, had *caused* the revolution. Two kinds of answers typically follow. First, the qualified yes: these technological media were necessary but not sufficient, they provided new capacities for organization that previous revolutions did not possess. Second, the concerted no: the technologies are important, but the necessary and sufficient *cause* of the revolution was "the people." No one (except Biz Stone and Mark Zuckerberg) believes that these tools actually cause revolutions.[2]

Both answers miss the mark, but they nonetheless point to one of those well-worn paths of argument. On the one hand there are technologies that create new relationships, new capacities, or re-arrange existing relationships of knowledge and power. On the other hand, there are the reassuringly familiar collectivities—like "the people" or "the public" or "the community." Sometimes information technologies are invoked as a threat to older forms of collective life; other times, especially in response to inflated claims about the power of those technologies, they are seen as irrelevant to the power of known collectives. Do information technologies connect existing collectivities or do they generate the conditions of possibility for new collectivities—maybe even new *kinds* of collectivity?

Over the last couple of decades many observers, both scholars and journalists, have clearly sensed that there is a problem here. The problem is an unsolved one, if the proliferation of recent terminology is any indication: *network societies, virtual communities, digital culture, cyber-cultures, social media, social software, digital natives, online communities, crowd-sourcing, crowd-funding, organizational networks, networked publics,* and so on.

Each of these terms conjugates an apparently straightforward technological thing with an apparently straightforward collective of some kind. But the result is apparently not straightforward. Instead, each one poses anew an opposition between emergent technology and stable collectives, strengthening the idea that the two are of different orders. In some cases, these terms are optimistic propositions that older kinds of collectivities can be intensified or expanded; in other cases (e.g. *digital divides, information plantations*), the conjugations point to more pessimistic conclusions.

Lurking behind such terms and debates is a much more general question. Contemporary information technology brings into relief a long-standing tension about the constitution of large-scale collectivities: namely, do they actually exist in any meaningful sense before they are constituted? Or are they artifacts of their technological intermediation? This tension between "natural" forms of community and mediation – particularly technological mediation – is one of the oldest stories that moderns tell about themselves. These collectivities need not know themselves (the way "the people" is sometimes said to); they may not even know they exist until they are shown to themselves through the operations of knowledge making and technology.

In this issue of LIMN, we asked contributors to address the problem head-on, and to consider the *nature of representing and intervening in collectives.* To pull apart claims about technology and collective kinds, we engaged not only scholars of the present, but also of the past.

IN THE 1890S, IN EUROPE AND AMERICA, a new kind of collective became an object of analysis: the crowd. The most famous diagnostician of crowds, Gustave Le Bon, constructed this concept out of a concern about civilization and its discontents: the discovery of the unconscious; the new urban realities of density, electric light, and public transport; and the eminently Victorian interest in the primitive within.

What Le Bon and others recognized was not just that people sometimes gathered together in a particular way, but that this way of gathering was tied to a particular moment in history, to a set of technologies and environmental changes and to hypothesized features of human behavior. The "crowd" was not just a horde or a mob, and it certainly was not polite society or a community. But it was new, and it was something that needed to be studied.

Fears about the crowd gave way within a few decades to increasingly sophisticated

1. Steve Lohr, "The Age of Big Data," *New York Times,* Sunday Review Section, Feb. 11th 2012. Curiously, the article is illustrated by the work of Chad Hagen, who creates "fictional data visualizations" that use no data. Chris Anderson, "The End of Theory," *Wired* 16(7).

2. And Wael Ghonim, whose memoir *Revolution 2.0: A Memoir* (Houghton Mifflin Harcourt, 2012) has garnered the most attention of this sort, though a similar kind of opposition is repeated in nearly every discussion of the Arab spring.

talk about "mass society" and the values and dangers of propaganda. Similar diagnoses—from the high cultural narratives of the Frankfurt School to the handbooks of propaganda, or the strategies of the new mass medium of radio—accompanied this new collective kind. The tension was also visible in the rise of a form of market capitalism that relied on anonymity—a mode of asocial, or anonymous, sociality that would eventually become a familiar problem for marketing, demographic research and national welfare. New kinds of collectivities are linked in obscure ways to the technologies that might make relations among people real, or visible, and sometimes both.

A similar story can be told about all of the heterogonous collective kinds that feature in our world: the public, the people, the population, the nation, society, the community (both the 19th century primitive community of ethnology and 20th century voluntaristic ones of communitarianism), the demographic segment, the network, and so on (see the infographic on the following pages). All of them have to some degree been 'naturalized' through the varieties of cultural practices that take them for granted: the design of government; the collection of information about people, including their behaviors and biology; and the attempt to use them as heuristics for the control of large groups of people.

Such a question is likely more familiar to historians than it is to those who claim expertise about new technologies. For instance, historians of statistics like Ian Hacking, Ted Porter, Alain Derosières or Mary Poovey have very clearly described the direct role of statistics in constituting "society" and "populations" in the 19th century. The technical characteristics of statistics coupled with national infrastructures of censuses, public health and policing called these new kinds of collectives – now taken for granted – into being. Sophisticated means of representing collectives, such as statistics, enable new forms of management, governance and intervention. They ultimately create a seemingly clear-cut concrete kind—a collective that people can occupy, analyze and ultimately govern.

Reflecting on the historical production of collective kinds can help orient and generate questions about these new phenomena—and some of the pieces in this issue provide that framing. In his piece, Alain Desrosières explains how debates about statistics after 1968 in France raised the question: "Do statistics have politics?" It pitted the "leftist" correspondence analysis (a technique used most famously in Bourdieu's *Distinction*) against the "rightist" neo-classical statistical thinking. As he points out, many of the same "leftist" approaches are now at the heart of datamining and profiling projects in "big data." Similarly, Rebecca Lemov's piece reflects on the kind of collective implied by the everywhere-and-nowhere device of the focus group. It emerged at the height of cold-war mass society and reflected mass society's desires back to it, through the artificial creation of representative individuals (who Joan Didion archly referred to as those "twenty people who lived in or near Cincinnati"). Both correspondence analysis and focus groups "map the collective" in different ways—either through a clever statistical technique that integrates the aggregate with the idiosyncratic or in the case of focus groups, by creating a space in which the idiosyncratic is allowed to stand for the aggregate in a way that is simultaneously convincing and absurd.

One of the most obvious collectives aggressively produced and represented in increasingly sophisticated ways is the political body of representative democracies—alternately figured as the people, the citizenry, the public sphere, the voting public etc. Daniel Kreiss and Maria Vidart outline for us what happens when social media collides with these classic collective kinds. The authors pose a double question of control: can social media be used to control voters and campaigners during an election, and conversely can social media itself be controlled? What kind of unruly new collective does it represent and what will be its effects on the established practices of mobilizing voters and winning campaigns?

The language of "crowds"—crowdsourcing, crowdfunding, the wisdom of crowds—has become one of the dominant modes of figuring the collective at the heart of new information technologies today. It is not the same crowd of Le Bon, though a comparison would no doubt prove fascinating. In their contributions, Alek Felstiner (a labor lawyer) and Roma Jhaveri (a former director at crowdfunder Kiva) present theoretical and practical accounts that explain clearly what these new techniques do well and what challenges or shortcomings they face. Lilly Irani shows us the detailed workings of Amazon Mechanical Turk—one of the most successful of the crowdsourcing endeavors. In her portrait she shows how AMT both solves problems that require human labor—the kinds of things computers still can't do—at the same time that it creates a new problem of management.

Tarleton Gillespie takes us inside (or as close as we can get to) an algorithm: the one powering "Twitter Trends." Big data is rarely interesting as such—rather it takes on significance in the moment when it is used to display a collective to itself, whether as a visualization of something or, as in the case of Twitter Trends, as a claim about some movement or trend of a collective rather than an individual nature. The question of whether such an algorithm can be wrong is not straightforward. Indeed, can one feel strongly—much less be right or wrong—about a *collective* without first finding a way to show that collective what it is?

Nick Seaver's piece also lays open the workings of big data, in his case the technique of "collaborative filtering" at the heart of software like Netflix and Amazon recommendation systems. Collaborative filtering reveals just how central—and how unquestioned—the notion of individual preference has become, and how it is being programmed into the heart of the tools we use.

Because it is so easy to look directly to social media and the Internet when asking about things like crowds and collectives, shifting the focus into different environments can reveal things overlooked. Natasha Schüll's contribution points us to the "touch-point collectives" of casino machine gambling. At the forefront of consumer data gathering, the closed world of casino redesign detects, constructs, and caters to specific collectives. Paying attention to these practices can diagnose larger concerns about data, privacy, consumer behavior and the control exercised by the corporations who own the data.

Similarly, Emmanuel Didier shifts our gaze to that of the police—specifically those in the Real Time Crime Center (RTCC) of the New York Police Department. NYC's police have gained notoriety for their use of statistics, and in particular for "compstat" which now routinely figures as a kind of artificial detective in crime dramas like *The Wire.* Didier shows not only how the RTCC works with data as a live stream, but also how it serves to create a form of police protection more suitable for Wall Street than other New Yorkers. Like Desrosières, Didier shows how "data mining" serves certain

political purposes and not others.

Chris Csikszentmihályi steps back even further to look at how engineering education is related to the kinds of technologies and problem solving that exist today. Engineering was for most of the 20th century the province of the engineers on the inside of the universities, the defense industry or the government. But with the advent of the Internet, and especially of Free and Open Source software in the 1980s, that dominance has begun to wane—today there are collectives of amateur engineers growing everywhere, and not beholden to the demands of mainstream engineering. Csikszentmihályi shows some of what such alternative engineering collectives might achieve.

Finally, the very emblem of resistance to the creation of new collective kinds is anonymity. From the anonymous Federalist papers of an 18th century public sphere, to the presumption of anonymity in markets, to the anonymous subjects of propaganda, the un-named and un-nameable are powerful figures of critique and danger in nearly every figuration of a collective. Gabriella Coleman puts the contemporary hacker collective Anonymous on display—both to show how and where they operate, in the technically specific domain of the Internet Relay Chat network, but also to show us how her own involvement as an ethnographer (and not a journalist) buys her membership (or not) in this collective.

The collection of articles in this issue shows the depth and diversity of perspectives that can interrupt conventional accounts of the phenomena of crowds and clouds. There are (new?) collectives of people and (new?) collections of data about which we actually know very little, and there is too often a demand to speak in haste, to claim expertise on the basis of familiarity, and to rely too easily on concepts such as "society" or "community" that should also be placed in question. The race for novelty in world of information technology should be a clear occasion for pause in the world of thought; and so it is here...

CHRISTOPHER M. KELTY
MARCH 2012

ZU DEN BLAUEN FLASCHEN: The scene portrayed in this painting is from Hof zur Blauen Flasche in Vienna, one of the first coffee houses in Europe when it was founded in 1686. ARTIST UNKNOWN, CIRCA 1900.

1700

1800

A B C E

A THE PEOPLE

The figure of The People is most clearly associated with Jean-Jacques Rousseau. Emerging in the age of Revolutions in response to new experiments with government, it figures the individual as a citizen bound by a contract with other citizens and through which a general will emerges. Rousseau's citizens are central to any story of democratic revolution, the rule of law or the cultivation of civic virtue, and they are concurrent with crucial innovations in the practical techniques of government—from the science of voting to the economic theories of the Physiocrats. The political tract, the assembly and the protest in the street are the techniques most dramatically pioneered in this era when the problem was, as Condorcet put it, to bring all of France into conversation with itself.

B THE PUBLIC (SPHERE)

In the Late 18th century, the public was a figure addressed in physical places like coffeehouses, and new media spaces such as Addison and Steele's 18th century periodical *The Spectator.* Most clearly tied to the salon, the *tischgesellschaft,* the broadsheet, and the pamphlet, the public turns individuals into a collective that pays attention, debates, argues and addresses each other, often anonymously. The Public was powerful because it represented legitimacy outside of authority, one which governments ignored at their peril. It was only as modern media grew into channels of money and power that the public sphere would become "public opinion" and the "re-feudalization" of the populace—as Habermas diagnosed it—would take place.

C THE POPULATION

Only in the early 19th century did the idea of the "population" as an entity with both structure and history emerge as an object of analysis in response to the hunger, disease and crowding of the Industrial Revolution. Most famously associated with Thomas Malthus' work *An Essay on the Principle of Population*, it was first and foremost a problem of governance: how to balance the uncontrollable geometric increase of the poor underclasses with the brute arithmetic facts of food production? Measuring the population would become a core obsession of the new science of statistics, the institution of the census, and the activities of public health reformers like Edwin Chadwick and criminal diagnosticians like normal curve inventor and astronomer Adolphe Quetelet.

D SOCIETY

Closely related to the concept of the population and arguably the most diffuse kind of collective, "society" emerges clearly as a real object of intervention only in the mid to late 19th century. In part, its recognizability is tied to the ill effects of the Industrial Revolution—worker accidents, effects of the poor laws, public health—rectified by new technique of social insurance and welfare. The "social" is clearly separate from (both inside and outside of) the nation or the citizen. In John Stuart Mill's famous analysis in *On Liberty*, social conformism is indicted as a tyranny at least as dangerous as the repressions of monarchies. It is in the rise of the anthropology and sociology—especially the works of Durkheim that the concept becomes a precise diagnostic of collective life.

E THE (ETHINIC) NATION

The rise of the nation-state is one of the most studied collective formations in history. Identified with the powerful mechanisms of patriotism, standardization and territorial military power, it has arguably been the most successful collective formation. Psychologically, nations tied ethnicity to moral duty; practically they incorporated people by unifying and standardizing languages and by taking advantage of the print, post and communication revolutions of the 19th century. Nations in their 19th century form, emerged in response to the revolutions and counter-revolutions of the period, and consolidated themselves through international economic institutions and military power. Nationalism would eventually become the pathology *par excellence* of the 20th century.

F THE PRIMITIVE COMMUNITY

In contrast to the nations, societies, the people or the population, the idea of "community" gains scientific salience with the rise of ethnology, archaeology, anthropology, and folklore in the mid-19th century. The distinction of Ferdinand Tönnies between *Gemeinschaft* and *Gesellschaft*, or that Henry Maine separating "status" and "contract" societies are among the most famous; primitive or natural communities were saturated with affect and sedimented in rigid class distinction as opposed to the societies and civilizations of Europe breaking free from tradition. Assimilating populations into the framework of nationality was a political concern at the same time that the "salvage" of primitive culture as a record of our past was in vogue among scientists.

COMMUNE: Black Bear Ranch friends and family. PRESS PHOTO FOR THE MOVIE *COMMUNE*. PHOTO BY JOCK STURGES, 1980.

1900   1920   1940   1960   1970   1980   1990   2000

D F   G   H   I   J K   L

G CROWD
The figure of the crowd at the end of the late 19th century was a hybridization of certain aspects of modern society and the primitive community. Modern civilization with its urban density, rail transport and globalizing markets was combined with the discovery of the unconscious, and the concern about the primitive, bestial drives that remain locked inside modern men. Crowds represented the uncontrollable eruption of these atavisms, exacerbated by the density of people, the ease of communication and the fragility of psyches battered by modern society. Le Bon and Freud would both fasten on the crowd as figure of the unconscious, while sociologists like Georg Simmel and Gabriel Tarde would attempt to make sense of the figure of the stranger, or the "mimicry" that led to a mob.

H MASS SOCIETY
Propaganda, nationalism, the new broadcast medium of radio, and the success of market capitalism are inextricably tied in the notion of "mass society" that emerges in the early 20th century. Often most clearly associated with facism, it was also the core object of analysis of the early communications and public relations scholars. It was also the object of critique of the Frankfurt School philosophers Adorno and Horkheimer, the "distracted viewer" of Walter Benjamin, and later of Hannah Arendt as well. As with the crowd, the notion of mass society referred to a manipulable and vulnerable unconscious as an object of control, this time not as an uncontrolled horde, but as an object of propagandistic direction and influence.

I THE DEMOGRAPHIC
Segmentation, pie charts, graphs and maps, target marketing, focus groups, surveys and opinion polls. All these techniques emerged in as mass society came to be measured and represented in new ways. In economic terms the "demographic" implicitly invoked a segmented consumer market—an expansion of capitalist productivity and a mode of representing tastes and desires in manageable, marketable ways. In political terms it allowed the periodic and constant measurement of opinion that political scientists would make into the object of a new science of democracy. Vance Packard would indict the advertisers of the 1950s in *The Hidden Persuaders*, while C. Wright Mills warned of the effects on political life in *The Power Elite*.

J (SOCIAL) NETWORK
Befitting the integrated national collective crisscrossed by telecommunications wires, postal systems, roads and railways, the notion of a collective as a network became most attractive in the 1960s. Perhaps the most iconic representatives are Stanley Milgram's "six degrees of separation" and Mark Granovetter's "strength of weak ties." The image of the network diagram representing the nodes and ties of a collective steadily grew into an object of both social and scientific fascination, promising (not for the first time) to give a precise mathematical definition to collective life.

K THE VOLUNTARISTIC COMMUNITY
Unlike the primitive community of the late 19th century, the community of the "communitarians" (in political theory) or the participatory democracy of the New Left and "New Communalists" (tuning in, turning on and dropping out), the 1960s and 1970s saw an efflorescence of forms and theories of intentional affiliation. Though the roots (especially in America) go back to the earliest socialist experiments, the new wave of community emphasized consciousness via psychotropics, the computer, and the commune; like the public sphere of yore, it experimented via the new media of 'zines, comics and video; from the primitive community of the 19th century ethnologists it took the spiritual and animistic as guides—but less for the coherence of a collective than for the expression of an authentic self.

L MULTICULTURALISM AND DIASPORA
The most recent transformations in notions of collective identity grew out of the critiques of nationalism, colonialism, mass society and public opinion, and in relation to the movements of liberation of the 1960s. Concepts of "hybrid identities" "hyphenated subjects," "diasporic communities," and "multiculturalism" took shape in cultural and aesthetic movements, educational curricula reform, and new attempts at recognizing and incorporating diversity into collective life. Collective life under multiculturalism affirms a cultural ethnicity without a national or biological determinism—and as such associates itself with the "soft power" of culture, aesthetics, tourism and self-expression.

THE WEAKNESS OF CROWDS
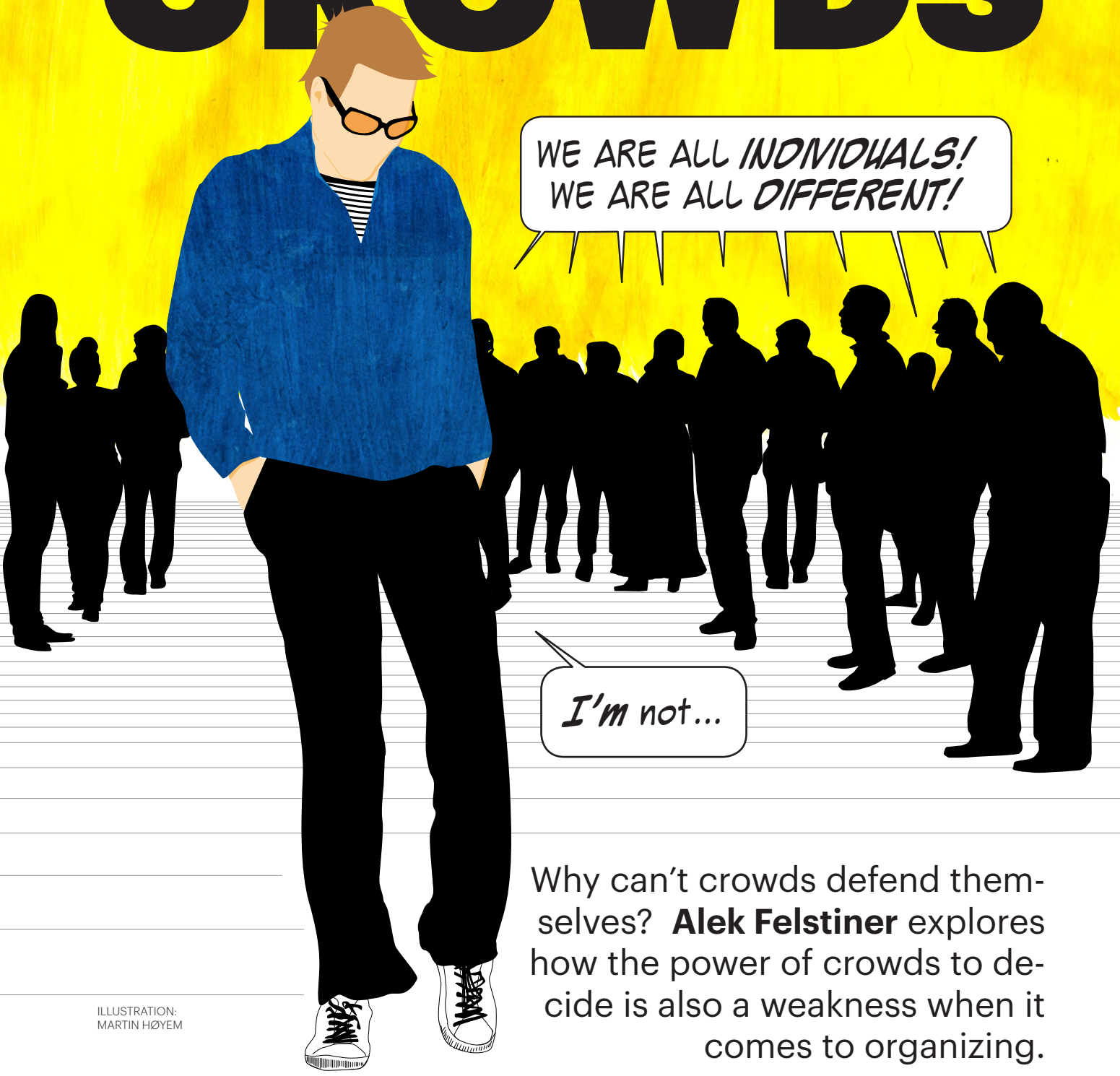
WE ARE ALL *INDIVIDUALS!* WE ARE ALL *DIFFERENT!*

I'm not...

Why can't crowds defend themselves? **Alek Felstiner** explores how the power of crowds to decide is also a weakness when it comes to organizing.

"IN THE FUTURE, ___ will be done by the crowd." This sentiment has become ubiquitous among those engaged in the study or promotion of technologies that exploit large, networked groups of people. Though I'm an employment lawyer, not a Silicon Valley entrepreneur or tech pundit, I confess to having expressed my enduring faith in the crowd on a number of occasions. It's quite seductive to imagine a future in which we solve all our problems through connecting with one another, aggregating our knowledge and our energy. Naturally, a futurist crowd-driven utopia also appeals to those who would look to the crowd as a substitute for – or an extension of – existing methods of production. Crowds can process and generate enormous volumes of data in sophisticated ways, with alarming speed, and at quite low cost. As such, it has become common to view crowds as an exploitable labor pool.

Commentators describe the phenomenal capacity of crowds in various ways, including "wisdom" (Surowiecki, 2004), "wealth" (Benkler, 2006), and "surplus" (Shirky, 2010). It boils down to the simple and perhaps self-evident premise that, in the right circumstances, we can get something more or better by sending "the crowd" to do what we would otherwise trust to the state, private enterprise, or individual agents. Yet I find myself wondering whether we spend too much time focusing on the strength of crowds, their ability to cohere, innovate, collaborate, and self-police. Crowds can be incredibly effective in building encyclopedias, developing products, processing data, constructing communities, entertaining themselves and others, and – in some cases – influencing powerful institutions. But we must be careful to avoid confusing the output of a crowd project with the individuals and groups that bring it into being. I would suggest that the very attributes that make crowd members so effective at *producing* also discourage them from engaging in collective action to protect and empower themselves.

A few years ago I began to explore the phenomenon of paid crowdsourcing, in which a group of workers accept and perform tasks on a web-based platform in exchange for compensation from whoever posted the task. I approached it as a legal and industrial curiosity, something that bore little resemblance to any existing labor model. By way of a brief legal detour, our labor laws were designed and enacted with the 19th and 20th century industrial employment relationship in mind. In crowd labor, there is no physical jobsite. Firms and workers connect on software platforms through various independent vendors who exert substantial control over the relationship. The technology of crowd production enables instantaneous contracting and compensation, personalized workspaces, and – most unusually – the integration of work into other online activity, such as gaming or social networking. Instead of the typical one-to-many relationship between employer and employees, crowd work platforms tend to encourage many-to-many relationships, each lasting only a short time.

I went about making the case that such activity should be classified as work, and that those who perform it should enjoy the protection of employment and labor law (Felstiner 2011). I was driven by my belief that employment laws should display internal consistency, and should, where possible, avoid falling hopelessly behind the emerging online work models that represent the future of our networked information economy. Those were my scholarly reasons, but I also had practical reasons to advocate for the employment rights of crowd workers. In a nutshell, I have more confidence in their ability to vindicate individual statutory rights than in their ability to exercise collective rights, or even to engage in collective action in any way that would accomplish a fundamental change in their working conditions.

Why should this be? If crowds are so collaborative, and generative, and influential, shouldn't they be able to use that power to improve their own individual and collective position? I believe they can, but they face structural obstacles created and reified by the way crowds are valued, cultivated, and regulated.

WHAT MAKES A CROWD VALUABLE?

In *The Wisdom of Crowds,* James Surowiecki identified three factors that make a crowd "wise": diversity of opinion, independence, and decentralization (Surowiecki, 2004). In this case, "wisdom" refers to the ability of the crowd to outperform experts in solving various kinds of problems. At the core of Surowiecki's argument lies the presumption that crowds become wiser the more their members behave as individuals. If you need a problem solved, don't get an expert or build a bureaucracy. Instead, use a wise crowd to aggregate individual judgments and actions into some whole that will be better than the sum of its parts. You get the all the benefits while avoiding the costs and distortions of centralized authority, and the pitfalls of groupthink, bandwagoning, and information cascades. In Surowiecki's ideal crowd, no one is too influenced by anyone else, too similar, or too connected. Those concerned about the potential for solidarity within the crowd should begin to see some red flags here.

Wisdom, or problem solving, is not the only thing we want from crowds. Where productivity is paramount, the wisdom of crowds takes a back seat to their volume and efficiency. Volume allows crowds to produce and process more information, more quickly. Where crowd members are sufficiently independent and diverse, per Surowiecki, a large crowd may also function to smooth out statistical anomalies. However, a crucial element of the voluminous and productive crowd, one that cannot be overstated, is fungibility. In order for a crowd to function as a crowd, rather than a large and unwieldy remote workforce, its members must be interchangeable. Fungibility plays the most prominent role in microtask crowd labor models, such as Amazon's Mechanical Turk, where large projects are broken down into many small pieces to be performed by any member of a sufficiently large crowd. But a certain amount of fungibility is present in

every crowd endeavor, in the sense that productive crowds do not and cannot depend on the contributions of any particular member.

Fungibility also makes crowd participation globally accessible, which certainly benefits workers seeking to earn a living through the crowd. In some cases, crowd labor offers a unique way for those in rural or developing areas to participate directly in emerging information economies, after only minimal training, and to be compensated well by their local standard. But it is important to remember that, while virtually anyone can participate, this is true because, under the demands of the crowd labor model, virtually anyone will do.

Some of the most ardent crowd evangelists also look beyond problem solving and productivity to the social, political, and cultural value of crowds. Yochai Benkler's *The Wealth of Networks* (2006) and Clay Shirky's *Here Comes Everybody* (2008) have shaped the modern canonical view of the crowd as a model of democratic participation, an avenue for unfettered cultural expression, and an alternative to the stagnating, institutionalized imperative to consume instead of create. People like Benkler and Shirky seem to value crowds primarily for their potential to transform our social and political life.

a well-designed crowd space can use technologies of various kinds to lower marginal costs such that free riding and collective action problems will shrink or dissipate. This permits crowd members to create rather than consume, to act outside the market, and to democratically structure their environment free from the influence of established hierarchical institutions.

I describe the various ways to "value" the crowd because crowds are less spontaneous than they may appear, and are often (if not always) the product of cultivation, and subject to regulation in some form. Multiplayer video games and crowd work platforms are examples of crowds deliberately cultivated by some preceding actor, usually with an interest in the crowd's growth and efficacy (however measured). But even crowds that seem to cohere or coalesce on their own will have members cultivating and recruiting others, and usually making some effort to self-regulate. The way we value crowds, our perception of their purpose, informs and guides how we cultivate and regulate them. If we want a "wise" crowd, we will endeavor to institute some version of Suroweicki's three criteria. If we want an efficient, on-demand, scalable workforce, or if we simply view such a workforce as an ideal opportunity for people

form of direct democracy, we will be emulating Benkler's vision of a wealthy network. Suppose it works. What happens then?

CONCEPTIONS OF AGENCY

Frequently, those advocating for the value of crowds – however conceived – are essentially appealing to our notion of agency. In the platonic crowd, membership enables each individual member, and by extension the collective, to maximize agency. Crowd members have the option to participate rather than passively receive, and the option to create rather than consume. They can communicate outside regimented media channels. They can sidestep institutions and existing markets.

This is one conception of agency, focused on the actions of crowd members in relation to the "output" of the crowd, and — perhaps — the production process. What about the ability of crowd members to determine the terms of membership? Or the future of a project, or the future of the crowd itself? Shirky characterizes successful crowd communities in terms of "promises" made to potential and existing members, "tools" used to coordinate the crowd, and the "bargains" or expectations of each member as a result of joining (Shirky 2008). How much control can individuals exert over these elements, particularly the "bargain?" Theoretically, these elements should be as susceptible to crowd influence as the crowd's ultimate "product," and we might even expect the crowd to determine these elements by necessity as it forms. Yet the things that make a crowd effective can also operate as impediments to collective action.

First, there is the problem of exit. Dissatisfied crowd members, like the consumers described in A.O. Hirschman's *Exit, Voice, and Loyalty,* are faced with the choice of discontinuing their participation or attempting to improve the crowd by speaking out alongside (or on behalf of) others (Hirschman 1970). Many crowd models expect and depend

> ## Crowds are less spontaneous than they may appear, and are often (if not always) the product of cultivation, and subject to regulation in some form.

Non-networked groups of people often fall prey to collective action problems, in which all members would benefit from some action but the cost to each individual is too great for him or her to undertake it alone. Similarly, the cultivators of public goods and creative spaces must usually endure some number free riders. Crowds too may suffer from these complications, but

with unfavorable alternatives, we will attempt to recreate the crowd dynamics characterized by Mechanical Turk and other microtask crowdsourcing platforms. And, if we are cultivating crowds in order to encourage non-market forms of cultural expression, privilege creation over consumption, or confront established political institutions with a legitimate and effective

on turnover, especially those focused on production. When everyone is fungible, and the numbers are so large, why would a crowd member choose voice over exit? The truly discontented, those who would rather leave than continue without some improvement, will look at the sheer size of the available pool, and the interchangeability of its constituents, and recognize the futility of becoming a squeaky wheel. Instead, they will just leave.

Second, there is the individual isolation and the dearth of existing institutional organizations. These represent the flipsides of the independence and decentralization so prized by Surowiecki. Collective action does not strictly require interpersonal proximity or available organizational infrastructure, but they can certainly help, and few collective efforts succeed without them. Moreover, when we consciously or unconsciously design a crowd to be "wise," to avoid the pitfalls of groupthink, bandwagoning, and information cascades, we risk throwing the baby out with the bathwater. After all, preserving independence to avoid groupthink also probably means foreclosing on the possibility of a strong consensus informed by shared deliberations, which is the backbone of an enduring collective effort. Similarly, solidarity represents the other side of the bandwagon effect, in the sense that people feel secure and supported in taking a stand when they know that others are doing the same. If we cultivate and regulate crowds to ensure that the decisions of one member have little impact on another, the potential for solidarity and consensus can disappear.

Finally, the typical crowd environment may ease collective action problems in some cases while exacerbating them in others. For example, crowd workers could employ collective action to improve their jobs and protect their interests by organizing themselves to accept no tasks below a certain wage floor, or by establishing mutual aid organizations (essentially unions) to train employees, provide certain benefits, and advocate on behalf of clickworkers to industry and government. Unfortunately, on a work platform such as Mechanical Turk, the volume and fungibility of the labor pool makes a wage floor campaign essentially fruitless and prohibitively risky. There will always be someone else available, equally qualified, and willing to work below the wage floor, because a global labor pool is always available and the requirements to join it remain quite low. These elements are built into the labor model, so crowd workers cannot create and leverage artificial scarcity. A mutual aid organization is slightly more plausible, especially on platforms that already encourage and make space for collaboration. Those could become solidarity spaces, though the efficacy of such organizations might end up roughly proportional to the percentage of the crowd they could legitimately claim to represent.

I am speaking here in broad and fairly abstract terms. Certain smaller crowds, or influential subgroups within crowds, will doubtless prove quite effective at expressing their collective will and achieving results. And conversely, some of those who establish crowd spaces and cultivate crowds may decide to foster collective action instead of chasing wisdom or productivity. They may decide that the value of an empowered collective outweighs the need to beat the experts and underbid the outsourcing subcontractors. However, I expect such projects will prove few and far between, given the strong financial incentive to do otherwise. The market for crowd production will continue to pursue the platonic ideal of an infallible, hyper-efficient, diverse, independent, decentralized crowd. When this happens we must recognize that the same things that make a crowd strong can also make it weak. □

ALEK FELSTINER *is an attorney at the U.S. Department of Labor. The views expressed in this article are the author's personal views and do not purport to represent the official views of the Department.*

WORKS CITED

Benkler, Yochai. 2006. *The Wealth of Networks,* New Haven: Yale University Press.

Felstiner, Alek. 2011. "Working the Crowd: Employment and Labor Law in the Crowdsourcing Industry," *Berkeley Journal of Employment & Labor Law,* 32(1), 143-204.

Hirschman, A.O. 1970. *Exit, Voice and Loyalty.* Cambridge, MA: Harvard University Press.

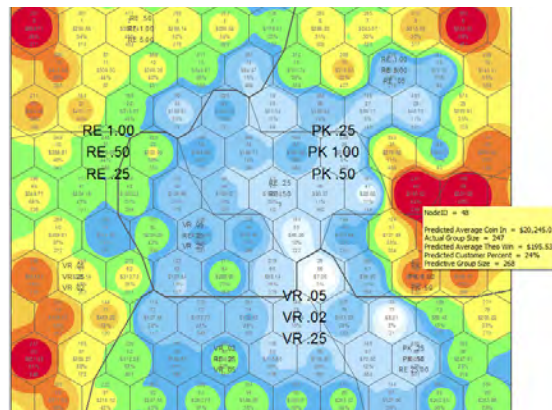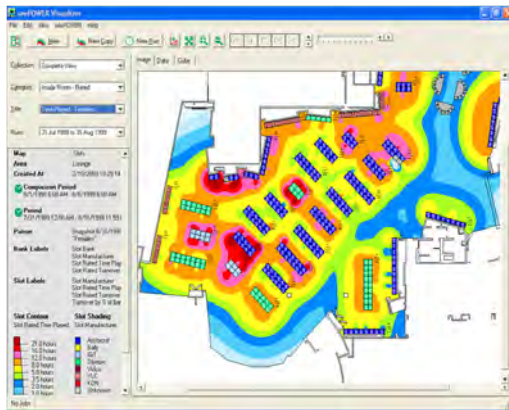Surowiecki, James. 2004. *The Wisdom of Crowds,* New York: Doubleday/Anchor.

The Touch-point Collective

Women under thirty and retired men might have surprisingly similar tastes for gambling. **Natasha Dow Schüll** explains how casinos have created a new kind of crowd.

Historically, casinos have been eager adopters of technologies that help them to gather knowledge about their customers. The knowledge-gathering repertoire of the modern casino has shifted from telephone surveys, focus groups, and rudimentary datasets to complex feats of reconnaissance and analysis enabled by player tracking systems, data visualization tools, and behavioral intelligence software suites. Many surveillance techniques first applied in casinos were only later adapted to other domains—airports, financial trading floors, shopping malls, banks, and government agencies.[1] "Knowledge is power and perhaps nowhere is this more evident than in the gaming industry," observed a trade journalist in 1999, before Internet corporations like Google, Amazon, and Facebook had become famous for their innovations in

Nearly 70 percent of casino patrons in the United States participate in so-called loyalty programs, using player cards to gamble rather than coins, paper money, or tickets. While their participation grants them redeemable points based on the volume of their play, it grants casinos a wealth of information. Casino player tracking systems, inspired by airline and credit card reward programs in the mid-1980s, record the value of each bet gamblers make, their wins and losses, the rate at which they push slot machine buttons, and what drinks and meals they purchase. Tracked gamblers are treated less as individual subjects than as "dividuals" in the Deleuzian sense—collections of traits, habits, and preferences that casinos can systematically compare to those of others in order to identify distinct customer niches (Deleuze 1992; Andrejevic 2007). Harrah's, a franchise that tracks

consumer monitoring and marketing (Parets 1999:19). "The unique advantage of our industry," notes gambling executive Lars Klander of TechResults, "is that we have hundreds of touch points a week or month, thousands per year, so you've got a profuseness of data that you don't have in other areas." Out of this data casinos draw the contours of customer groups that we might think of as "touch-point collectives." An examination of the ways that casinos detect, intervene in, and shape these groups offers insight into the workings of such collectives in other contemporary domains.

players seamlessly from coast to coast by pooling information from its national chain of properties into a single centralized database, parses its market into ninety different segments and addresses each with a unique marketing scheme.[2]

In the past, strategic adjustments to casino layout, game selections, or marketing campaigns were based either on trial and error or on projective modeling techniques such as "stochastic migration," in which theoretical individuals created from samplings of actual behavior (typically recorded by note-taking observers) were

CONTOURING SOFT-WARE
◄ "Each hexagon is a group of players. Close hexagons are more alike than distant hexagons."

◄◄ Casino floor pictured by seePOWER Visualizer.

GAMBLER INTELLI-GENCE
◄◄ The illustration on the previous page is from an advertisement in which Bally Business Intelligence promises that with their help you will "see your players, your games, your floor, and your business like never before."

---

1. Casinos used biometric systems, for instance, well ahead of law-enforcement agencies, airport security, and businesses (Schwartz 2003: 216–17). They also used non-obvious relationship awareness (NORA) software used to discover cheating collusion in casinos before it was adopted by Homeland Security to investigate connections between terrorist suspects (see Kaplan 2010). Yet the exchange goes both ways: emerging systems of server-based gambling, for instance, adopt the same cryptography system as the government's National Security Agency.

2. Harrah's Total Rewards website (https://Harrah's.com/TotalRewards/TotalRewards, accessed November 2010). The Total Rewards program began as Total Gold in 1998 and changed names in 2000 with the addition of "player tiers." By 2004, 30 million cardholders were signed up in the program, and today the number stands at 40 million.
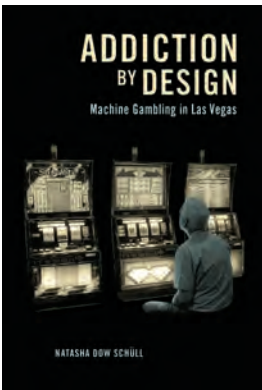
This article draws on a chapter from the article author's book, *Addiction by Design: Machine Gambling in Las Vegas,* 2012, Princeton University Press.

followed through virtual simulations of an environment before and after a proposed design change—moving a bank of machines a few feet left, for instance, or widening the entryway to a gaming area. While the challenge for such techniques was to gather enough of the right information on existing customer behavior to forecast the effects of a particular change, the challenge for tracking techniques is different: How to extract meaningful insight from a continuous stream of "live data" that is overwhelming in its volume and detail? "We've got 20,000 [behavioral] models per second, streaming off the floor," remarks Javier Saenz, vice president of network systems at IGT.

The challenge has given rise to behavioral intelligence software suites like Compudigm's seePOWER, a technology that promises to help casinos cope with "atomic level data across the enterprise in real time." SeePOWER transforms massive amounts of touch-point data into colorful heat maps that represent collective "tendencies and preferences" in and over time, as a company press release puts it. A

Every evening at the same time, female patrons under thirty were moving from one side of a popular bank of slot machines to the other ...

representative gave a demonstration of the software's abilities on her laptop computer during the 2007 Global Gaming Expo in Las Vegas. She pulled up a map titled "Time Played—Females," picturing data for women players on one casino slot floor over a twenty-four-hour period. Dark red clots of color appeared around the machines they had played the longest, encircled by bands of progressively "cooler" shades around those they had played less—magenta for sixteen hours, pink for twelve hours, and so on. (See illustration on previous page.) In a time-lapse animation, these shaded contours undulated as gambling activity waxed and waned at different machines. An animation of the same area over five consecutive days revealed a curious pattern. Every evening at the same time, female patrons under thirty were moving from one side of a popular bank of slot machines to the other, while males over fifty were taking their original seats. Upon further investi-

gation it was discovered that the men were exiting a nearby showroom at the close of a revue performance and pestering the young women. Casino managers responded by creating "a whole new protective area for those women" containing the machines they liked to play, and sent out a direct mailing to promote the new slot shelter. Previous profit levels not only returned but were quickly surpassed.

The casino's data cloud, when animated and queried, had rendered visible the fleeting, real-time contours of a behavioral group whose constituents, seated at individual play terminals and immersed in the solitary activity of play, were likely unaware of their kinship. Casino managers attempted to profit from the proclivities of this touch-point collective by carving out a physical space for its members and formally inviting them to gather there—not to socialize, but to continue to interact with their own game screens. Although the players were affiliated by age, gender, game preference, and ultimately a common gathering site, the collective they formed was "virtual" in the sense that it took shape and subsequently became meaningful through casino data analysis and visualization software rather than through self-selection, voluntary participation, or shared experience.

The collective-making abilities of Compudigm's software extend beyond the physical confines of the casino, generating "outside maps" that complement "inside maps" by further illuminating the behavior of particular groups. "Let's say we want to see the profitability of females fifty-five and older. Who are these ladies? Where do they live? How can we target them better?" The representative showed an animated map of an unidentified city, titled "ground floor, little old ladies, carded play time." As the clock in the upper left-hand corner spun, the city flared and pulsed with color, registering the home addresses of older women gamblers as they began and ended sessions of machine play on the ground floor of one casino over the course of a day. In the wee hours, small circles of color dotted the landscape, with red centers indicating the neighborhoods most heavily populated by current onsite players. Starting at 8 a.m., the center of the map dramatically blossomed outward into a bright red flower, reaching maximum size at 11 a.m. and shrinking back in the evening; across the

city, discrete pockets of "little old ladies" continued to gamble throughout the night. Armed with this knowledge, the casino was in a position to tailor its offerings to the play schedules and affinities of the market segment in question.

Complementing seePOWER's ability to detect and track the habits of particular behavioral collectives through time is its capacity to comparatively visualize the respective profitability of diverse demographic segments. The company spokesperson presented one such map in which cumulative levels of spending for different slot types appeared as color-coded circles (red for high spending, shades of orange and gold for moderate, green and blue for low). Over these ponds of color a neat honeycomb lattice was superimposed, each of its hexagons containing labels for the group characteristics corresponding to the spending intensities beneath (e.g. "middle income, ethnically diverse," "retired, single income," and the like). (See illustration on previous page.) The map's colorful profitability contours, overlaid with the geometric grid of market segmentation, were presented as a visual tool for revealing hidden group likenesses and distinctions that could guide casino managers' marketing interventions.

As in the case of online venues like Amazon.com, individuals' consumer behavior in casinos, recorded in a common data cloud and refracted through statistical analysis, becomes the basis for group classifications of which they may not be aware but in which they continue to participate—and in sometimes more robustly as a result of the customized product marketing that follows. The contours of touch-point collectives are honed through an iterative process of data differentiation and marketing response that tends toward a telescoping of group tastes.

Server-based gaming, in which game content can be instantly downloaded from an online server or "jukebox," is an example of a marketing response that promotes this kind of telescoping. On the one hand, the technology allows casinos to anticipate the tastes of predefined cultural, ethnic, and demographic groups; game screen font sizes, for instance, can be increased in anticipation of the arrival of elderly patrons, while slot machines can be virtually "reskinned" for special events like Cinco de Mayo or Chinese New Year (Macomber and Student 2007: 28). On the other hand, it allows casinos to flexibly react to unanticipated preference groups in real time, switching out game configurations (i.e theme, denomina-

tion, payout rate) to match player predilections as they emerge and shift. As tracked touch-point data is continuously uploaded to the casino's data cloud and analyzed in the aggregate for collective patterns, digital game content is continuously downloaded from the game cloud such that the casino becomes "dynamically responsive" to the affective and behavioral contingencies of its player markets (Green 2007:34). Knowledge and intervention are intimately linked in the creation of touch-point collectives. □

NATASHA DOW SCHÜLL *teaches at the Massachusetts Institute of Technology in the Program in Science, Technology, and Society.*

WORKS CITED

Andrejevic, Mark. 2007. *iSpy: Surveillance and Power in the Interactive Era.* Lawrence: University Press of Kansas.

Deleuze, Gilles. 1992. "Postscript on the Society of Control." *October* 59: 3–8.

Green, Marian. 2007. "Station Casinos Carefully Rolls Out Guaranteed Play Option to Video Poker Crowd." *Slot Manager,* November/December.

Kaplan, Michael. 2010. "How Vegas Security Drives Surveillance Tech Everywhere." *Popular Mechanics,* January 1, http://www.popularmechanics.com/technology/how-to/computer-security/4341499, accessed August 2009.

Macomber, Dean, and R. Student. 2007. "Floor of the Future I." *Global Gaming Business* 6 (11).

Parets, Robyn Taylor. 1999. "Advances in Linked Gaming Technology." *International Gaming and Wagering Business* (Special Issue for World Gaming Congress and Expo), September: 19-20.

Schwartz, David G. 2003. *Suburban Xanadu: The Casino Resort on the Las Vegas Strip and Beyond.* New York: Routledge.

Crowd Funding and its Challenges

The concept of crowd funding isn't a new one. People have long pooled money together to help a neighbor in need, participated in a shared business investment, or contributed money towards a larger cause. The introduction of online crowd funding platforms such as Kiva.org offers a new spin on this established practice.

BY **ROMA JHAVERI**

SYSTEMS LIKE KIVA allow like-minded people to pool their funds through formalized web platforms rather than through informal family connections or community associations. While these platforms create valuable efficiency gains around the collection and transfer of funds from one set of people to another, they also raise questions about the role of the platform. In particular, what – if anything – is the role of the platform in managing the crowd which it serves?

This is a question that affects the core of Kiva, the world's first personal microlending website. Kiva's mission is to connect people through lending to alleviate poverty. In six years of operation, Kiva has facilitated close to $300 million in microloans from lenders in 218 countries to borrowers in 60 countries. The scale of such an accomplishment is made more remarkable by the fact that Kiva has consistently maintained an internal staff of less than one hundred employees. This very scale and diversity of Kiva's reach, combined with its relatively small staff, poses unique challenges. Because it is impractical for Kiva to be intimately familiar with each end-recipient of Kiva financing, it is neither possible nor appropriate for Kiva to attempt online promotion of borrowers' who may seem most in need or most worthy of financing. Kiva lacks the deep knowledge of local contexts

and borrowers circumstances necessary to make such judgments. Instead, Kiva positions itself as an open marketplace where all parties can act autonomously according to their own values. By being an open marketplace, Kiva tries to let lenders, field partners, and potential recipients make the value judgments of making and taking loans.

A more detailed understanding of Kiva's workings helps to illuminate this position. The Kiva ecosystem consists of multiple types of crowds, rather than just one. There are the obvious crowds: the over 650,000 lenders who have banded together to make loans and the 750,000 borrowers who have received loans and made repayments. But there are also the smaller crowds. One-hundred and forty-three Microfinance Institutions (referred to as "Field Partners") from around the world choose the local borrowers that they post on Kiva's website. These partners raise funds, administer the loans to those borrowers, and aid in the transfer of funds. More than two-hundred Kiva volunteers edit and translate loans entered by Field Partner staff. Another four hundred volunteers have served as Kiva Fellows, working directly with Field Partners to help increase the ecosystem's impact where loans are managed and dispersed. Managing this number of indi-

PREVIOUS PAGE
Lorenda Millamena wears a hat, one of the many bamboo products she sells, in her sales shed in Sibalom, Antique Province, the Philippines. Lorenda, a 55-year-old mother of ten, has used microloans from Kiva partner Ahon Sa Hirap, Inc. as working capital for her business selling charcoal and raw and finished bamboo products. PHOTO: HODAG.

THIS PAGE
Lucinda Alonzo at her market stall in Bugasong, Antique Province, the Philippines. Lucinda has used microloans from Kiva partner Ahon Sa Hirap, Inc. to start her business selling dried saltfish and shrimp paste. PHOTO: HODAG.

viduals becomes increasingly difficult when you consider that each distinct type of crowd has unique needs of the Kiva platform – needs that are sometimes in tension.

Field Partners, for example, would like to optimize the amount of zero-capital financing they receive from Kiva for their borrowers while minimizing the amount of work done to receive these funds. As organizations striving towards sustainability, this is a logical cost-benefit approach for them to take. It follows that Field Partners have an incentive to post those borrowers whom they believe are most likely to quickly attract funding on Kiva. They determine who these borrowers are by process of trial and error. For instance, if a butcher funded more quickly than a beekeeper in the past, the Field Partner would more likely post butchers than beekeepers in the future.

Borrowers, by contrast, need to receive financing via Kiva's site while still maintaining personal dignity and privacy. To respect this need, Kiva must strike a balance between providing enough information about a borrower so that lenders can make funding decisions, but not so much that the Kiva website intrudes on borrowers' personal lives. To enable lenders to make funding decisions, Kiva represents each borrower on its website with a photo (often taken at the borrower's place of business), a location, and a description of the borrower's life circumstances. To ensure that borrower privacy and safety is prioritized, Kiva offers Field Partners the option of abstracting certain details about the borrowers' lives. For example, the Field Partner can choose to disclose only a borrower's country of residence rather than their specific town.

Where borrowers might seek privacy, Kiva lenders often want to feel that they are making as much of a difference as the platform permits. When borrowers and Field Partners have abstracted their descriptions, some lenders have taken issue, claiming that abstractions limit their ability to fully imagine the living circumstances of borrowers. In a crowd funding platform, the funding crowd determines who

is most deserving of funding; these lenders often push Kiva to offer information they feel enables greater loan impact. These platforms are marketplaces of fixed resources in which the lending community is empowered to vote on who should and should not receive funding through their individual lending decisions, as well as collective mobilizations through the platform.

Kiva's lending community recognizes this power and has wielded it to create change on and through the platform. Cockfighting loans became one example of values in tension, negotiated between Field Partners and lenders through the platform. Cockfighting is legal in Peru and Field Partners have posted loan requests for cockfighting businesses on Kiva in the past. Some Kiva lenders who saw these requests understandably objected to this practice on the grounds of animal cruelty. Kiva's policy, however, is to permit fundraising for activities that are legal in the borrower's country and are consistent with major UN conventions. Kiva relies on law and codified convention to avoid the blurry lines that could come to delineate which borrowers could raise funds on its platform – lines that could become subject to the whims of Kiva's staff members. The argument for preventing animal cruelty, for example, could be stretched to prohibit fishermen from appearing on Kiva's platform.

Because Kiva's stance aligns with major UN conventions and local Peruvian law, Kiva permitted the cockfighting loans to raise funds on Kiva's website despite their moral challenges. Somewhat unexpectedly, the lending community's negative reaction ultimately had an impact on the Field Partners posting the loans. Recently, Field Partners have backed away from posting cockfighting loans. Their explanation has been that it was never their intention to culturally offend.

Kiva's role in balancing all of these differing needs is to provide as much information as possible to its crowds to inform their decisions. By showing Field Partners that demand for funding loans is greatest in December, Kiva encourages them to maximize their loan postings to

Kiva for December. Similarly, providing lenders with information about borrowers and Field Partners enables lenders to find borrowers that align with their personal social and philanthropic interests. Kiva's newest attempt to inform lenders is the Social Performance Center on Kiva's website, which calls out the social performance strengths of Kiva's Field Partners. These strengths include, for example, "Client Voice," "Anti-Poverty Focus," and "Entrepreneurial Support." Through these designations, lenders can focus their lending towards Field Partners who emphasize support borrowers in particular ways.

There are limits, however, to the guidance that Kiva will provide. The open marketplace philosophy that Kiva continues to embrace is sometimes criticized for its failure to mobilize lenders around particular borrowers. Some perceive it as an obstinate refusal to provide guidance about which borrowers are most deserving of funding. But given the many differing parties and interests involved – most of which are only touched on here – treating borrowers equally is the best available option. To prescribe solutions around which individuals are most deserving of our community's resources is to flirt with the idea that Kiva possesses the "correct" worldview. This is an impossible idea for a platform that – due to the varied crowds that comprise it – actually brings different worldviews into exchange. □

ROMA JHAVERI *served as Product Manager and then Director of Product Management at Kiva.org from 2007-2012.*

Can an Algorithm be Wrong?

How do we know if we are where it's at? **Tarleton Gillespie** explores the controversy over Twitter Trends and the algorithmic 'censorship' of #occupywallstreet.

ILLUSTRATION: MARTIN HØYEM

THROUGHOUT the Occupy Wall Street protests, participants and supporters used Twitter (among other tools) to coordinate, debate, and publicize their efforts. But amidst the enthusiasm a concern surfaced: even as the protests were gaining strength and media coverage, and talk of the movement on Twitter was surging, the term was not "Trending." A simple list of ten terms provided by Twitter on their homepage, Twitter Trends digests the 250 million tweets sent every day and indexes the most vigorously discussed terms at that moment, either globally or for a user's chosen country or city. Yet, even in the cities where protests were happening, including New York, when tweets using the term #occupywallstreet seem to spike, the term did not Trend. Some suggested that Twitter was deliberately dropping the term from its list, and in doing so, preventing it from reaching a wider audience.

The charge of censorship is a revealing one. It suggests, first, that many are deeply invested in the Twitter network as a political tool, and that some worry that Twitter's interests might be aligned with the financial and political status quo they hope

to challenge. But it reveals something else about the importance and the opacity of the algorithm that drives the identification of Trends. To suggest that the best or only explanation of #occupywallstreet's absence is that Twitter "censored" it implies that Trends is otherwise an accurate barometer of the public discussion. For some, this glitch could only mean deliberate human intervention into what should be a smoothly-running machine.

The workings of these algorithms are political, an important terrain upon which political battles about visibility are being fought (Grimmelmann 2009). Much like taking over the privately owned Zuccotti Park in Manhattan in order to stage a public protest, more and more of our online public discourse is taking place on private communication platforms like Twitter. These providers offer complex algorithms to manage, curate, and organize these massive networks. But there is a tension between what we understand these algorithms to be, what we need them to be, and what they in fact are. We do not have a sufficient vocabulary for assessing the intervention of these algorithms. We're not adept at appreciating what it takes to design a tool like Trends – one that appears to effortlessly identify what's going on, yet also makes distinct and motivated choices. We don't have a language for the unexpected associations algorithms make, beyond the intention (or even comprehension) of their designers (Ananny 2011). Most importantly, we have not fully recognized how these algorithms attempt to produce representations of the

wants or concerns of the public, and as such, run into the classic problem of political representation: who claims to know the mind of the public, and how do they claim to know it?

THE CONTROVERSY ABOUT TWITTER TRENDS and #occupywallstreet may be, by itself, a tempest in a teacup. But even on its face, censorship was a difficult charge to dismiss. Of course, some may have made or repeated this charge as a casual observation, as a gesture of belonging, as an expression of frustration about their political movement not being heard, or as a deep-seated anxiety about the nefarious intentions of corporations. But for those who leveled this critique with more care and conviction, the first question was, what exactly does Twitter measure when it identifies these Trending terms?

Twitter has repeatedly stated that their Trends algorithm is not a simple measure of volume (i.e. the most used terms), though the minimalist presentation of the list may suggest otherwise to some users. Some of the comments charging censorship, whether of #occupywallstreet or Justin Bieber, suggest that this list is often taken as a straightforward and unproblematic measure of pop-

standing into their everyday use of Twitter.

Until 2010, Twitter had not provided much, or any, explanation of its algorithm. What Trends represented was offered as self-evident. When similar charges of censorship were raised around #wikileaks, Twitter responded on their blog,[1] and there laid out some general criteria. (These criteria were corroborated and further explained by a Twitter engineer, in the comment thread of one of the more widely-read critiques.[2])

Twitter explains that Trends is designed to identify topics that are enjoying a surge, not just rising above the normal chatter, but doing so in a particular way. Part of the evaluation includes: Is the use of the term spiking, i.e. accelerating rapidly, or is its growth more gradual? Are the users densely interconnected into a single cluster, or does the term span multiple clusters? Are the tweets unique content, or mostly retweets of the same post? Is this the first time the term has Trended? (If not, the threshold to Trend again is higher.) So this list, though automatically calculated in real time, is also the result of the careful implementation of Twitter's judgments as to what should couns as a "trend."

described in the blog, and these criteria undoubtedly change over time. For instance, one substantial revision occurred in May 2010 when Twitter announced it was removing Justin Bieber from the Trending Topics list.[3] Further, their explanation does not say enough: for instance, how something like "clusters" are defined and measured remains opaque.

Twitter may not be able to say much more about how Trends works. Revealing the "secret sauce" of their algorithm in greater detail risks helping those who would game the system. Everyone from spammers to marketers to activists to 4chan tricksters to narcissists might want to optimize their tweets and hashtags so as to Trend. This opacity makes the Trends results, and their criteria, deeply and fundamentally open to interpretation and suspicion.

The Trends algorithm and the data it evaluates are also the property of Twitter. Sites like Trendistic[4] can use the Twitter API and the corpus of public tweets to conduct more exhaustive analyses of the volume and velocity of term. But they cannot access private tweets, and they cannot know or take into account what counts as a cluster of users. Despite what Twitter is willing to make known, any effort to discover the Trends criteria can only amount to sophisticated guesswork.

Trends is also part and parcel of Twitter's financial aspirations as a for-profit business. Whether or not Trends is an accurate or a deeply flawed measure of vital topics of discussion, it is

> One substantial revision of the Trends algorithm occurred in May 2010 when Twitter announced it was removing Justin Bieber from the Trending Topics list.

ularity. Though a few critics went to great lengths to reverse engineer the Trends results, to uncover the criteria that enliven them (Lotan 2011), most users certainly vary in their comprehension of what Trends measures and how, and may not always incorporate that under-

Of course, these are just the publicized descriptions of what Trends looks for, and they do not offer a definitive (or satisfying) answer to critics. Trends measures something both more precise and more obscure. There are likely more—and more specific—criteria than those

also a means to entice and retain users. This does not necessarily mean that it must squelch volatile topics like #occupywallstreet. Twitter has trumpeted its role in the democratic uprisings of recent years, pitching itself as a vital tool for political foment. If Trends is meant to high-

1. http://blog.twitter.com/2010/12/to-trend-or-not-to-trend.html
2. http://studentactivism.net/2010/12/05/wikileaks-twitter-3/#comment-11619
3. http://mashable.com/2010/05/14/twitter-improves-trending-topic-algorithm-bye-bye-bieber/
4. http://trendistic.indextank.com/

light terms that will support Twitter's self-proclaimed relevance, #occupywallstreet seems to fit the bill. But this political vitality must be balanced against the interests of other users, of advertisers, of regulators (Gillespie, 2010). Most of all, it is not as if these competing interests can be simply weighed and settled. This same algorithm must serve the desires of its users, all the while drawing new ones. The degree to which it can do all that is the most pressing criteria for Twitter.

As such, Trends is both an index of what is said and an advertisement to read further. Trends promises to bring new readers to a topic; this accounts for much of why the Occupy activists care if and when it appears. The act of measuring the phenomena, then, also changes it: Trends is both feedback and feedback loop, because something that does Trend may be discussed further.

Trends provides an algorithmic gloss of our aggregate social data practices that can always be read/misread as censorship. If #occupywallstreet is not trending, that could mean any of the following: (a) it is being deliberately censored (b) it is actually less popular than one might think (c) it is very popular but consistently so, not a spike (d) it is popular and spiking, but not in a way the algorithm is designed to measure (e) it is popular and spiking, but not as much as some pop culture phenomena that has crowded it off the list (f) it is popular and important, but not as popular as the pop culture phenomena that have been strategically gamed onto the list (g) it has not Trended because it has not Trended, thereby not enjoying the amplification Trends itself offers However, we are unable to know for certain. Not only are the criteria opaque and the archive proprietary, we also have no benchmark against which to compare the results. Trends measures activity on Twitter, and it is only implicitly indicative of broader public concerns.

TWITTER TRENDS IS JUST ONE OF MANY information practices that claim to know or represent the will of the people: public opinion polls, census surveys, man-on-the-street interviews, voting mechanisms. Each employs a specific technique to assess public opinion or activity, in order to make the public will legible. With Trends, Twitter is making a claim: that surging terms, measured in a specific way, are indicative of topics of the most interest and import. But this claim is, for Twitter, caught between competing desires: reporting back what Twitter users care about most versus drawing new users into new conversations. Users too are caught between competing desires: wanting to know something accurate about the public beyond them, and wanting to be rendered visible as a part of that public. With both of these tensions at work, the politics of the artifact, i.e. the workings of the Trends algorithm, become just one piece of a greater puzzle: the politics of representation.

What's different here is that Trends promises a mathematical and an exhaustive analysis of what is being talked about, while presenting it as automatically generated and self-evident facts about the discussion. It claims to know the public through an algorithmic assessment of their complete traces, which is different than the professional judgment of a newspaper editor, the zeitgeist insight of the fashion trendspotter, or the statistical approximations of a census taker. Further, these indices are rendered in an instant and built immediately back into the service itself.

Twitter Trends is only one such tool. Search engines, while promising to provide a logical set of results in response to a query, are in fact algorithms designed to take a range of criteria into account so as to serve up results that satisfy not just the user, but the aims of the provider, their understanding of relevance or newsworthiness or public import, and the particular demands of their business model (Granka 2010). When users of Apple's Siri iPhone tool begin to speculate that its cool, measured voice is withholding information about abortion clinics, or worse, sending users towards alternatives preferred by conservatives, they are in fact questioning the algorithmic product of the various search mechanisms that Siri consults.[5]

Beyond search, we are surrounded by algorithmic tools that offer to help us navigate online platforms and social networks, based not on what we want, but on what all of their users do. When Facebook, YouTube, or Digg offer to mathematically and in real time report what is "most popular" or "liked" or "most viewed" or "best selling" or "most commented" or "highest rated," they are curating a list whose legitimacy is built on the promise that it has not been curated, that it is the product of aggregate user activity itself. When Amazon recommends a book based on matching your purchases to those of its other customers, or Demand Media commissions news based on aggregate search queries (Anderson 2011), their accuracy and relevance depend on the promise of an algorithmic calculation paired with the massive, even exhaustive, corpus of the traces we all leave.

We might, then, pursue the question of the algorithm's politics further. The Trends algorithm does have criteria built in: criteria that help produce the particular Trends results we see, criteria that are more complex and opaque than some users take them to be, criteria that could have produced the absence of the term #occupywallstreet that critics noted. But further, the criteria that animate the Trends algorithm also presume a shape and character to the public they intend to measure, and in doing so, help to construct publics in that image.

A term that has trended before has a higher threshold before it can trend again. The implication is that the algorithm prefers novelty

5. http://www.rawstory.com/rs/2011/11/29/10-things-the-iphone-siri-will-help-you-get-instead-of-an-abortion/

in public discourse over phenomena with a longer shelf-life. This is a longstanding critique of broadcast journalism, reappearing in social media. Perhaps we could again make the case that this choice fosters a public more attuned to the "new" than to the discussion of persistent problems, to viral memes more than to slow-building political movements. Trends also measures the velocity of a term, but within Twitter's network; it does not compare this surge of interest with matching attention on Facebook, say, or on the blogosphere, or in traditional news coverage. So this public is understood to be platform specific, though we know that users participate in and manage overlapping networks of information and people. With a different commitment to understanding public discourse, one might design an algorithm that gives greater recognition to a topic surging across multiple platforms than one that only surges inside a single network. When Twitter restricts its attention to Twitter, though it is plain why they might want to do so, they put their finger on the scale of a debate about how political discourse does and should function online.

But here's a harder problem: Twitter takes into account whether a term is circulating within a pre-existing cluster of users (users who are already interconnected, following each other, regionally co-located, etc), or spans clusters. Trends presumes that topics are more important if they exceed clusters. The choice of how to know a

trend matters. Twitter could have designed Trends to weigh heavily a term that does not span clusters of users but instead powerfully ignites a single cluster of users. This kind of "trend" might spotlight issues and concerns discussed intensely by a small but engaged community. It certainly would have regarded #occupywallstreet more highly. Putting such terms in front of all Twitter users via the Trends list would intervene not by showing users what lots of people are talking about, but about what some are talking about hotly together. It might add issues to the public docket rather than rehearsing them. Choosing instead to value terms that span clusters is a choice, and a political one—more Habermas (1989) than Mouffe (2000) in its implicit theory of political dialogue, for example.

These algorithms produce not barometric readings but hieroglyphs. At once so clear and so opaque, they beg to be read as reliable measures of the public mind, as signs of "us." But the shape of the "us" on offer is by no means transparent. Social media tools like Twitter may be adept at mapping networks of people, if only because they provide the substrate within which these networks form and interact. Even if they cannot as easily capture the human networks that extend beyond their own services, they certainly can claim to have scrutinized the part that is rendered on and by their system. However, though they aspire to with algorithmic tools like Trends, they

may not be as adept at identifying or forging the publics that emerge from those networks. □

TARLETON GILLESPIE *is an associate professor in the Department of Communication and the Department of Information Science at Cornell University.*

WORKS CITED

Ananny, Mike. 2011. "The Curious Connection Between Apps for Gay Men and Sex Offenders." *The Atlantic,* April 14: http://www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-apps-for-gay-men-and-sex-offenders/237340/

Gillespie, Tarleton. 2010 "The Politics of 'Platforms.'" *New Media & Society* 12(3): 347-364: http://hdl.handle.net/1813/12774

Granka, Laura. 2010. "The Politics of Search: A Decade Retrospective." *The Information Society* 26(5), 364-374.

Grimmelmann, James. 2009 "The Google Dilemma" *New York Law School Law Review* 53(4): http://works.bepress.com/james_grimmelmann/19

Habermas, Jürgen. 1989. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society,* trans. Thomas Burger with Frederick Lawrence. Cambridge, MA: MIT Press.

Lotan, Gilad. 2011. "Data Reveals that Occupying Twitter Trending Topics is Harder Than it Looks." *Social Flow,* October 12: http://blog.socialflow.com/post/7120244374/data-reveals-that-occupying-twitter-trending-topics-is-harder-than-it-looks

Mouffe, Chantal. 2000. *The Democratic Paradox.* London: Verso.

Top Needs of Occupy Sites Across the Nation

Healthy Food @OccupyCleveland

Wool Socks @OccupyToronto

Sterno @OccupyRochester

People @Occupy_UNH

Brooms @OccupyNS

Cardboard & Markers #OccupyHarvard

Blankets @Occupy_Boston

Printed Fliers @Occupy_Syracuse

Photo Copier @OccupyBGM

Books @OccupyWallStNYC

Media Equipment @OccupyNewark

USB Mic @OccupySetonHall

Clothes @OccupyNJ

Power Inverter @OccupyPhilly

Gas Generators @OccupyKSt

Tarps @Occupy_DC

Bags @OccupyRestonVA

Sleeping Bags @OccupyRaleigh

Duct Tape @OccupyChc

Dry Rice & Beans @OccupyJaxBeach

Cinder Blocks #OccupyColumbus

Can Opener @OccupyToledo

Hand Warmers @OccupyFortWayne

Hot Food @OccupyDayton

Laundry Service @OccupyBtown

Coffee @OccupyChicago

Firewood @OccupyWinnipeg

Whole Grain Bread @OccupyMinnesota

Alcohol Swabs @OccupyDenver

Propane @OccupySLC

Winter Equipment #OccupyEdmonton

Whiteboard @OccupyCalgary

Help to Survive @OccupyMissoula

Kool-Aid & Crystal Light #OccupyMiami

Wooden Pallets #OccupyCHA

Charcoal @OccupyHsv

Lantern @OccupyMemphis

Water Jugs #OccupyMemphis

Rain Gear #OccupyNewOrleans

Ice @OccupyTulsa

Tents @OccupyHouston

Sleep @OccupyAustin

Intel i7 Laptop #OccupySanAntonio

Computers #OccupyDallas

Supplies #OccupyTucson

Propane @Occupy_Oklahoma

Canopies #OccupyPhoenix

Hub & 3G Network Cards @OccupyTijuana

Sugar & Creamer @OccupyRiverside

Rain Gear @OccupyOC

Tupperware Bins @OccupyLB

Water @OccupyLA

Warm Drinks @OccupyFresno

Gardening Supplies @OccupySF

Cookware @OccupyOakland

Megaphone @OccupyMills

Beanies @OccupySacto

WiFi Hotspot @OccupyOregon

Batteries #OccupyPortland

Maalox #Occupy_Seattle

People #OccupyVictoria

Sandbags @OccupySpokane

NOV 3ʳᵈ Emergency Meeting

Transcript from meeting at Occupy Wall Street. Observed and transcribed by Amira Pettus.

Deshaun: *Josh took my bag of clothes from the back and gave them away. He told me he didn't believe in possessions, and that nobody should have them. He came in and tried to take down our tents, and tried to move the Comfort store.*

Mae: *He is doing us like he did SIS [Shipping Inventory and Storage]. They have banned him. He gave away all our sleeping bags, and he keeps being incredibly disruptive. It is really cold, and we told him we did not think sending sleeping bags to Washington Square was a good idea. He got a truck and did it anyway. People were really cold that night.*

Josh: *You guys have no right to do what you are doing. I am doing what is best for Comfort and you guys should listen to me. I can do whatever I want, there are no rules here. You can't enforce anything.*

Ivan: *No you cannot do whatever you want, if you do that we will never get organized. You are being disrespectful of our community.*

Josh: *You guys aren't doing things right, and you need to listen to what I am telling you.*

Patrick: *You need to listen to us.*

Cristine: *Josh stole money I saw him.*

S.X.- Facilitator: *That isn't relevant*

Cristine: *Yes it is, he is a thief*

Hiro: *No, it doesn't matter*

Mae: *It doesn't matter. I think that Josh should never be allowed to work here, or welcome as a participant in our community's process.*

S.X.- Facilitator: *Is that a proposal?*

Mae: *Yes.*

S.X.- Facilitator: *The motion on the floor is to evict Josh from Comfort, and Comfort activities.*
All in favor? [All hands go up, with exception of J.]
All opposed? [J.]

S.X.- Facilitator *Josh, the community has spoken, please respect their decision. This meeting is adjourned.*
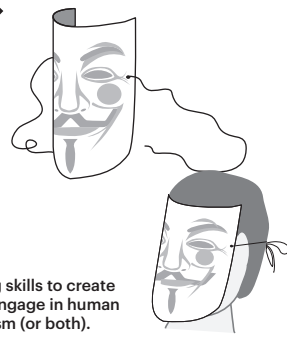
❶ Cut along the dotted line.

❷ Punch holes through the black dots next to the eyes. (You might want to make holes for the eyes as well.)

❸ Thread a string through the holes you just made and tie the mask to your head.

❹ Use insightful hacking skills to create fearsome pranks, or to engage in human rights technology activism (or both).

```
A1: everyone trusts you, so you're doing something right
A1: someone irl did say to me once
A1: 'oh you're talking to that biella again'
A1: 'shes SO a fed'
```

AM I ANONYMOUS?

Learning how Anonymous works means learning to be one. **Gabriella Coleman** narrates her experience of being in between worlds.

IT WAS DECEMBER 2010, and my plans were simple: finish my book manuscript on the politics of free and open-source software hacking and spend time with my family on an island off the coast of Washington State. That is, until Anonymous once again reared its head. While family members went hiking during the day and watched movies late into the night, I huddled over my laptop obsessed with Anonymous: a name and a cluster of ideals taken by different individuals and groups to organize distinct and often unrelated actions, from fearsome pranks to human rights technology activism.

Although by winter of 2008, individuals deployed various political demonstrations and activities under the banner of Anonymous (prior to this, the name was used almost exclusively to stage Internet pranks), it only fully entered public consciousness in December 2010. Unfolding before my eyes was a distributed denial of service (DDoS) campaign: #operationpayback. No doubt my research appeared rather lifeless to those around me; but what I was witnessing on Internet Relay Chat (IRC)—the central nervous system of so many geek and hacker interactions— was anything but boring. Normally home to lively, albeit quotidian and mundane conversation, scores of individuals populated the chat room #operationpayback, where actions were discussed and coordinated. At one point the channel ballooned to seven thousand participants and bots. Many were contributing to the DDoS campaign aimed directly at disabling the servers of Visa, Mastercard and PayPal. Julian Assange's organization Wikileaks, had just caused a major political firestorm by releasing 220 leaked confidential diplomatic cables, and these companies were targeted by #operationpayback for refusing to accept donations to Wikileaks.

For most of December I watched the blizzard of activity on AnonOps in silence, unsure how or when to intervene given the furiously fast pace of the conversations, spanning various topics, from the time-honored tradition of humorously taunting the FBI, to ethically dense deliberations on the DDoS as protest tactic. In early January, my silence came to end when a handful of Anons singled me out:

```
A1: Can anyone in here confirm bi-
    ella?
A2: i talked to her today but...
A3: you know her A2?
A2: if she would send me a DM on
    twitter, i could.
A3: "biella is away: I'm not here
    right now" and no @'s in any of
    7 channels...
A2: yes, if she's the biella from
    twitter, i talked to her before
A1: We may need to dispose of jour-
    nalists from here in just a bit.
A1: (Temporarily.)
A3: she can come back later
YOU HAVE BEEN KICKED BY A2: (hi bi-
    ella, could you DM me on twitter
    please? thanks!)

[I log back in, quite nervous]

BIELLA: hello A2 A1
```
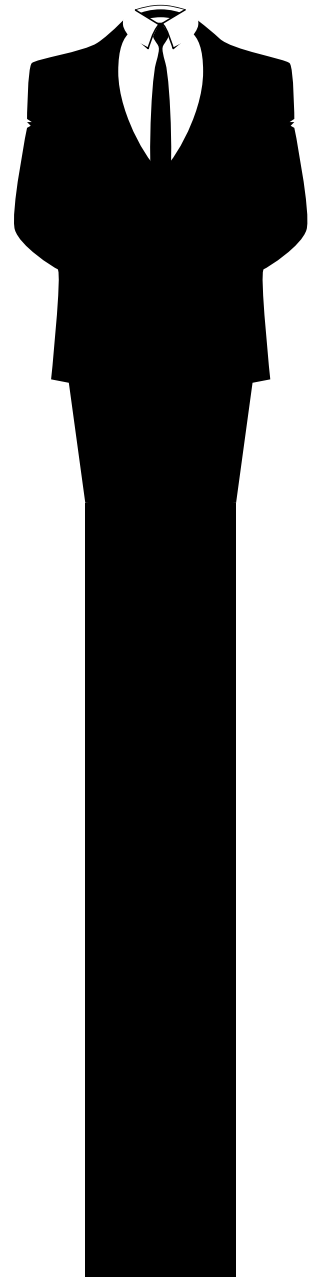
```
BIELLA: sorry about that i was away
   cooking
BIELLA: this is me
   [ . . .]
BIELLA: i have referred many re-
   porters here
BIELLA: and am writing/presenting
   on Anonymous
   [ . . .]
A2: Hi biella, apologies for the
   kick.
BIELLA: no it is ok
BIELLA: you gave fair warning :-)
   and i have been too too idle
BIELLA: more than i would like
A1: We're just usually very strict
   and sometimes a little paranoid
   of unidentified users in here.
```

It was a make-or-break moment. If these Anons had cast me in an unfavorable light (whether untrustworthy or a nuisance or both), it could have put an end to my research. These Anons not only seemed to be fine with my presence, some were keen to have me around. After this conversation, I chimed in more frequently, spending on average about five hours a day on IRC, roughly following six to twelve IRC channels at once, seven days a week. Following activity on AnonOps and a few other Anonymous networks has been simultaneously exhilarating and frustrating. Anonymous is clandestine and sprawling, with a constantly mutating labyrinthine architecture. In any moment there can be two to five active IRC networks, each populated by dozens of channels where Anons interact, sometimes seriously, sometimes playfully. Sometimes it is impossible to tell the two apart. Over the course of a mere fifteen minutes in a single chatroom, people might be joking about 'fapping' (aka masturbation), holding a serious discussion about the latest anti-piracy legislation under consideration in Congress, answering questions posed by a visiting reporter, launching virulent accusations against individuals, and greeting the visiting anthropologist. Take for instance, the conversation below, which reflects the multi-layered, multi-threaded, somewhat chaotic, and often quite playful character common to conversational life on IRC:

```
S: Three officers were also taken to
   hospital with serious injuries,
   according to police. Warsaw.
   [reporting on clashes in Poland
   on Independence day]
ANONREPORTERX: Will anonymous ever
   appoint any kind of leadership
   or known spokes people? Why/why
```

```
   not?
J: if there are no leaders, and the mass
   is not a leader either
J: who would have the capacity to 'ap-
   point leaders'
J: ?
ANONREPORTERX: I am asking.
S: True leaders speak for everyone.
ANONREPORTERY: let's do it anyway. it's
   Neil Young's bd. that would be a great
   present #anonspoxNYoung
M: Anonymous not longer is anonymous if
   it has an appointed leader..
J: anonreporterX, I know, and I'm trying
   to show you how, simply using common
   sense and logical reasoning
J: you can reason that there will never
   and can never be a leader
J: without having to even ask it
J: :)
K: ^
mode (+v biella) by S
J: a wild biella appears!
P: Oh snap
***x catches the wild biella
X: :p
BIELLA: :-)))
J: lol
BIELLA: a sleepy biella /me just wakin'
   up
BIELLA: glad to see this here [since a
   reporter channel had been down for
   awhile]
J: good morning then :P
[. . .]
M: Aloha!
X: lol
P: lol
X: aloha!
ANONREPORTERX: that seems unrealistic:
   1. Anonymous is already having to deal
   with defining who does and does not
   represent the movements' intentions
   (blac bloc in Oakland) and...
X: no
J: anonreporterX, where is it defined who
   does or does not represent Anonymous?
J: last time I checked, anyone trying to
   do so was talking out of his ass :P
```

As the conversation was unfolding, and prompted by AnonreporterX's trite question about leadership, I told one Anon that I would like to write an anthropological piece on journalist's obsession with leaders. During this private conversation, he followed with a question and comment of his own:

```
A8: about what?
BIELLA: how the medias desire to find a
   leader says more about a reporters re-
   lationship to their editor and certain
```

```
    obsessions within American culture
BIELLA: than anything else
A8: thats true
A8: I have yet to see a european other
    media obsess over leadership like us
    does
BIELLA: EXACTLY
A8: though uk tends to sensationalize it
    too
```

As might be obvious, much of my time with Anonymous is spent chatting on public channels, in back channels, and with single Anons and often without much aim; while I ask Anons targeted questions, I also go with the flow, doing as everyone else around seems to be doing..

The aimlessness is important, however, for it captures one of two important types of labor and interactivity valued by Anons. One is a form of charismatic sociality quite common on IRC where cleverness, cunning and playfulness garner attention and sometimes, even respect. The form of verbal interactivity and dexterity common to IRC is similar to a certain style of talk described as the "man of words" by the famed folklorist of African-American cultures Roger Abrahams. "A man of words is nothing" explains Abrahams, "unless he can, on the one hand, stitch together a startling piece of oratorical rhetoric, and on the other, capture the attention, the allegiance, and the admiration of the audience through his fluency, his strength of voice and his social maneuverability and psychological resilience." Abrahams differentiates between two categories of the man of words: one who displays stunningly crafted rhetorical flourishes in formal settings; the other, springs to life informally and spontaneously on the street corner, the yard, and especially over rum, speech characterized by playful, lewd, and more crass talk. Unsurprisingly, it is this latter type of verbal play and dueling common to IRC, although with some important differences, given the unique technological features of this technical space.

Despite the playful, sometimes brazen, and often boisterous atmosphere of laughter, pleasure, and verbal play common to IRC, Anonymous is still rather serious business, Which brings us to the second form of labor and interactivity crucial to gaining respect on the network. Anons (on AnonOps, among other Anonymous networks) acquire respect by engaging in *activist* interventions, some of them risky and illegal; there have been over two dozen arrests. By laboring toward collectively-defined political actions and by working on the infrastructure that supports this type of work (such as running an IRC server), individuals come to trust each other. One of the key operators and organizers of OpTunisia, which provided technology assistance to Tunisian activists in January 2011 and helped catalyze the string of Anonymous-led interventions in the Middle East region, explained this dynamic as follows:

```
BIELLA: but i am trying to figure
    out how it is that people come
    to start working with others and
    trusting each other
BIELLA: you seemed like a good
    person to ask as you have been
    around for a long time, know
    lots of folks, etc etc. it is
    just is so enigmatic and perhaps
    that is what it is
A: well i think either doing some-
    thing that gains you respect
    and in the process gets you
    'friends'
A: also if people help me i feel
    inclined to help in return
BIELLA: so what is an example of
    something you did that gained
    that respect (ofc keep it legal
    :-))
BIELLA: and also can you elaborate
    on the 'friends' bit
A: well i founded and coordinated
    op tunisia
BIELLA: ok, yep, i can see why that
    would gain respect ;-)
BIELLA: i did not know that
A: so i worked very hard for a
    while 4hrs sleep a night online
    20hrs a day
A: for 2ish weeks
BIELLA: and people started contrib-
    uting and you all felt prolly
    close as a result
A: yeah so up popped some individu-
    als - who are now 'famous' and
    said can we help and i worked
    with them
BIELLA: like hacker types you mean?
A: yeah ;)
```

If Anons accrue respect from a combination of charismatic sociality and especially work, what about me? I am not running an IRC server, nor do engage in political actions. Certainly, all the hours I have poured into IRC has been central to forging lines of communication and building trust among (at least some) Anons. I can hold my own on IRC and I rather like chatting on IRC, which may explain why I have chosen to study geek and hacker worlds: collective worlds that are inseparable, at some fundamental level, from this communicative architecture. But at a certain point, it became patently obvious that my research was rather more complicated than simply "hard chatting on IRC." I was also putting some labor into the collective pot. Indeed,

I hold the dubious distinction of teaching roughly two dozen reporters how to find Anonymous and how to get on IRC to interview them. For most of the winter and spring of 2011, I helped shuttle reporters onto the channel designated for them. I subjected myself to the mindless repetition of being interviewed over eighty times by journalists. I have answered the same questions over and over again in print, in TV and in film interviews. After a few months of doing this type of media-work—and it quickly came to feel like the drudgery associated with some forms of work —it became evident that I was gaining some access, respect, and trust via these appearances, many Anons peppering me with comments, reflections, praise, and critiques after they watched a news segment, read an opinion piece, or watched some public lecture.

My ethnographer's magic, to borrow a famous term coined long ago by Bronislaw Malinowski, may lie in how I handle myself in public lectures and the media: something I never expected when commencing this project. The work of ethnography is often about the private lives and thoughts of individuals or concerns public modes of interaction, not acting as the public face, in this case, of a faceless entity. I have earned some measure respect because I have worked assiduously to dispel myths. And I have had to literally engage in some cunning to do so, because so many journalists, especially in the United States and the UK, have been keen on slotting Anonymous in the role of raging hackers, led by a small cohort of leaders, or some other distortion.

In my many media appearances and talks, I state things that Anonymous themselves would not say (or would certainly put in different terms). Sometimes I just flat out contradict them. For instance, in the past, many Anons used to say "we are not hackers," a claim that became much harder to make once the hack-as-leaking operations took off in March 2011. I would explain: there *are* hackers but Anonymous is not simply composed of hackers. And sometimes, most significantly, I am silent; there is a lot I don't say or even currently put into written word.

As I recently explained to one sympathetic reporter in a lengthy interview on the ins and outs of studying Anonymous:

"There are things about Anonymous that I currently can't write about because I don't understand it well enough. You have to have some discretion because there are some back-room politics, and they need time to develop before you make a claim about it."

I also explained that I might be caught up in webs of duplicity myself:

"I'm aware that I am operating within webs of duplicity. While I've come to trust certain Anons and have more empathy than less, I'm also well aware that duplicity is the name of the game—misinformation and social engineering—and I'm being caught up in it myself," observed Coleman. *"But, if it was clear cut and transparent, it wouldn't be as effective politically."*[1]
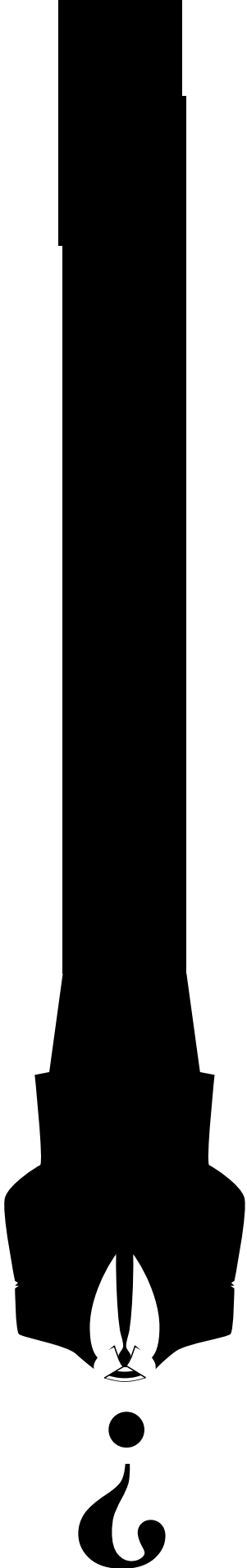
If Abraham's identified the man of words, a mode of talk also integral to communicative life on IRC, it might be best to describe myself as *the woman of measured words,* at least when I appear in the media or when I give a talk. Since I am hyper-aware Anons will critically asses, even at times dissect my statements, I am quite deliberate in what I say and don't say in public, as I know this will affect and shape my access to them. This does not mean I am simply cowered into silence. In fact, being blunt about certain issues—like acknowledging how I too may be the object of misinformation—has brought some measure of approval. But it is always a question of cunning and craft as to how, where, and when to make statements about Anonymous.

On IRC, like those around me, I often give way to the spontaneity of verbal play and meandering conversations. During my interaction with reporters, I take a distinct and measured stance. Most anons who pay attention to these things (many do not) witness these two sides, each performative in their own right, although requiring distinct forms cunning. Do these interactions—deliberate public media work and spontaneous socializing on IRC—make me Anonymous? □

GABRIELLA COLEMAN *is the Wolfe Chair in Scientific and Technological Literacy at McGill University.*

---

1   http://www.deathandtaxesmag.com/157192/digital-activism-from-anonymous-to-occupy-wall-street-a-conversation-with-gabriella-coleman/

▶ **OCCUPY SOURCING:** Occupy Wall Street rapidly re-created many of the functions of collective life—from libraries and kitchens to bureaucracies and markets. The working group structure is illustrated on the next page by Amira Pettus.

OCCUPY SOURCING

GENERAL ASSEMBLY

The People's Kitchen
As a symbol of solidarity, this image was created and used as a tattoo by the People's Kitchen.

Bookmobile
The #OWSBookmobile pickup tour was created to help rebuild the People's Library so that it may continue to be a source of knowledge, as well a place for freedom of expression.

Missed Connections
One of the poems from Craigslist's Missed Connections section to an Occupier.

Hoyt/Schermerhorn G

This weekend.
You had
an occupy wall street poster.
I had
a book.

Comfort
Minutes,like the one shown here, are used by the Occupation to document meetings and assemblies.

sleeping bags to Washington Square was a good idea. He got a truck and did it anyway. People were really cold that night.

J: You guys have no right to do what you are doing. I am doing what is best for Comfort and you guys should listen to me. I can do whatever I want, there are no rules here. You can't enforce anything.

I: No you cannot do whatever you want, if you do that we will never get organized. You are being disrespectful

Lego Librarian
Lego representations were made to show the changing events of Zuccotti Park.This one depicts the arrest of Scales, a librarian, during its eviction.

Crossing Occupations
A letter from Occupy Portland reaching out to Occupy Wall St. with the intent to share practices and an open line of communication.

Printers Guild
To make the movement more visible in physical spaces and not just virtual ones,Occuprint and the Screen Printers Guild ask that images like this one be made,distributed, and worn.

Size of Working Groups in Occupy Wall St. by Number of Members According to nycga.net

250 - 350 +
150-250
50 - 150
25-50
5 - 25

ENGINEERING COLLECTIVES

TECHNOLOGY FROM THE COOP

Engineers make the world, but not just as they please. **Chris Csikszentmihályi** recounts how engineers come to be part of one collective or another. ▶

Lieutenant Corporal Vaughn's reply to Sergeant David Linley (above) was one of dozens left on Linley's blog at betweenthebars.org.[1] Between The Bars is a blogging platform for the (other) 1% of Americans who are incarcerated, the vast majority of whom have no access to the Internet. Prisoners send handwritten letters to the site and, if they are not censored by the prison, they are published (and collaboratively transcribed) as a blog entry. Visitors to the blog can leave replies, which are then printed and sent back to the prisoner. Linley, a returned marine suffering from PTSD, wrote a few posts to Between the Bars, at first receiving only a few courteous responses from readers. About six months later, one of his fellow servicemen discovered his post and over the next few weeks more than a dozen veterans were letting him know how much he meant to them, sending care packages, and even visiting him in prison. Linley's case demonstrates how online media can help transform loose online social links into significant "in real world" support in times of need. In the case of Between The Bars (BtB), this was not a coincidence: it was designed by Charles DeTar, a researcher at the Center for Civic Media (C4) at MIT, specifically to help prisoners exercise online self-advocacy, an important prerequisite for collective action and social change.

Though DeTar is a PhD candidate at MIT, the Institute (where I taught for 10 years) is not a hotbed of technology development for progressive causes like prisoner's rights. Indeed, with most of its money coming from the US government (over 70% of its funding, most of that military) and massive corporations (nearly all the rest of the funding, from companies like BP and Bank of America), MIT largely embeds the needs of the most powerful in society into durable technologies. Nearly every contemporary engineering research institution is funded through similar models, and as a result the bulk of technologies entering the world ultimately reinforce the status quo. For example, technologies for prisons and law enforcement are a significant market for high technology research and development, while technologies for prisoners' rights and public defenders are not. Engineering education covers thermal dynamics and differential equations, but its funding structure also means that engineers must be taught

LETTERS FROM THE OTHER 1%
The vast majority of Americans who are incarcerated have no access to the Internet. betweenthebars.org is a blogging platform for this (other) 1% of the population: prisoners send handwritten letters to the web site and, if the letters are not censored by the prison, they are scanned, published, and collaboratively transcribed as a blog entry.

---

1. http://betweenthebars.org/blogs/129/william-d-linley-david?page=1

to work easily only in areas that support the most powerful entities in society. DeTar's work, then, is a form of activist technology, a stark contrast to the normative values of the institution in which he works. He is one of a growing number of technologists who bypass the normative structures of technologist education and professional identity by anchoring his work in a different dialog, the Free/Libre Open Source Software movement. F/LOSS is not an inherently progressive movement, but it does offer the largest, most powerful, and most sustained alternative to conventional technology education, development, and distribution. In addition, the free software movement has provided working models for new methods of Internet-enabled collective action that inspired Between the Bars and many other platforms for community collective action that we developed at C4.

ENGINEERING IDENTITIES

Engineering is socially regressive for several reasons, but perhaps first and foremost because the vast majority of engineering employers—government and corporations—expect their employees to help maintain the status quo. Schools like MIT, RPI, and Carnegie Mellon receive vast sums of government and corporate directed research, and entire areas of enquiry (Artificial Intelligence, Aeronautic and Astronautic Engineering) are primarily—in some cases exclusively—funded by commercial and military contracts. Undergraduates interested in these disciplines must either acquiesce to working on military research or (in many cases) regressive government or corporate research, or they must drop their vocation and find another major. For some time I wondered why my own research group, which focused on developing technologies for social justice and agonistic politics, received so many enquiries from sophomores and juniors in Aeronautical and Astronautical engineering; later I realized that many were uncomfortable with the nature of their other funded research opportunities. These students, attracted to

flight, were gradually realizing that advanced research in the field was rather about *fight.* No one explicitly told them they must toe this line; rather, faculty rehearse subtle narratives of professionalization, rationalizing why military funding is not only necessary but irrelevant to advanced research. Take for example the following dialog, nearly identical to the countless examples that I experienced at MIT, between an NPR interviewer and NYU computer science professor Peter Belhumeur:

NPR: Do you get government funding in part?
Peter Belhumeur: *Yes.*

NPR: From DARPA or one of those?
PB: *We just say the Department of Defense.*

NPR: Uh huh. You just say it. Do they say "This is what we're interested in?" or do they keep their cards close to their chest and say "We like this, here's some money.
PB: *They definitely say what they're interested in. I think they want to do face recognition and verification in the wild, in unconstrained environments. So, where the person in the photograph does not necessarily cooperate. And you can imagine why that sort of thing is important to them.*

NPR: So, especially in war as we have come to know it, in counter-insurgency operations and all that, it would be useful, as opposed to in the traditional battlefield where who cares who that guy is?
PB: *That's, that's, that's right. And one of the reasons that face recognition doesn't go away is because it's basically this passive biometric, and you can acquire the data at a distance.*

NPR: Does the ultimate application of this, the ways it could be used, ever give you pause?
PB: *Well, you know we think about it a lot, certainly within the group, and I don't think we're at the point at which these sorts of biometrics can essentially label people with perfect identity or any of that. And I think that there are interesting policy questions that surround*

that. And you know personally I'm on the side of this that less is better.

NPR: Uh huh.
PB: *But I think it's a really interesting scientific question.*

NPR: Will you reach a point though, as you get better and better and better in your scientific research some years hence where your predisposition to think that "less is more" comes to a head with "look how good we've gotten."
PB: *Yeah, I think at that point I'll stop. But there's no danger of that yet.*

NPR: Really? You've still got another 10 years of making it better?
PB: *There's still a lot of work to do.*

What can we learn from this dialog? First, a computer scientist should do work the DOD asks for, and should not make public the details about that work. Second, that work is a "really interesting scientific question" with "interesting policy questions," the former of which is the purview of the researcher but not the latter. Third, the engineer's personal feelings, explicitly at odds with the DOD ethos, are immaterial; he will work on the problem nonetheless. Fourth, he will keep working on the problem the DOD is paying him to solve until the point where his misgivings are realized and something awful has entered the world, at which point he might stop.

Many students accept and repeat these narratives, and learn to subsume their ideals or interests in directions that are militaristic or market-oriented shadows of what they had gone to school to study. Students interested in imaging are steered toward computer vision for weapons; ones interested in robots are directed into military drone research; others interested in environment or green energy are funneled into the interesting sounding MIT Energy Initiative. MITEI (pronounced "mighty" by its members) is funded primarily by BP, Schlumberger, and Halliburton, and seems bent on maintaining petrochemical hegemony). Engineering education

can be seen as the first in a series of filters within professional engineering that systematically remove individuals interested in challenging societal power, or remove the will to challenge from individuals.

A second filter is the professional identity of American engineers. Engineers learn that they cannot influence (and thus should not bother thinking about) the course of technologies. This is partly due to the tacit client relationship to power, but it is also part of a larger social and intellectual history described by Matthew Wisnioski in his dissertation, *Engineers and the Intellectual Crisis of Technology, 1957-1973* (Ph.D, Princeton, 2005). Wisnioski describes how through the 1960s an intellectual firestorm raged over how to think about technology. One faction argued that technology was a semiautonomous agent, able to drive history and change society, though it was not co-influenced by society or history. Any negative quality of a technology—what engineers call "unintended consequences"—derived from the natural tendencies of the technology. Somehow "through proper study technology could be managed" (even though it could not be influenced by society or history!). The second faction argued that technology was influenced by society, and indeed reflected the values of its builders. This more political view argued that if technology seemed to be running amok, it was a reflection of the priorities of the society behind it, and society itself should be changed. The majority of engineers adopted the former theory, of technological agency, which absolved engineers of responsibility for technology's negative effects but undermined the engineer's role as a creative, autonomous agent.

In choosing to limit their liability, engineers had to construct a complex, self-denying logic that dissimulated their own daily thinking, planning, choices, indeed their labor. Despite the fact that no engineer believes that technology is autonomous in the particular -- at the scale of their own daily work -- they nonetheless adopted a "zoomed out" view that erased the

Students ... learn to subsume their ideals or interests in directions that are militaristic or market-oriented shadows of what they had gone to school to study.

social aspects of their profession. Furthermore, while they might have personal identities, and might think that a particular kind of work is immoral or unethical, these factors matter less than the technoscientific "interestingness" of the problem. They are oracles of technology, taking orders from political agents (like the DOD or DARPA, Monsanto or Schlumberger) yet somehow purporting to remain apolitical themselves. Engineering education and professional identity doesn't so much inculcate ethics as systematically separate technical work from ethical thought and action. Ultimately, a professional engineer must subsume their own moral, political, and intellectual agency, channeling instead the interests of their clients.

I FREED MY SOFTWARE, SO I FREED MY MIND.

The origin story of the Free Software movement has been well described: generally it is said to have launched when a bearded and poorly socialized programmer named Richard Stallman, frustrated that copyright prevented him from fixing buggy commercial software later developed a nice bit of legal jiu-jitsu which enables software to be simultaneously copy-written yet forced into the public commons (Kelty 2008). Different historians concentrate on different aspects of this history, and certainly this legal coup is important, as was Stallman's development of the GNU compiler and operating system, and later Linus Torvalds' related work with Linux. What is often underdescribed is the actual mechanism of online collaboration, using various technically enabled tools and communication methods that help to coordinate a geographically distributed labor pool of heterogeneous individuals; many of whom have never met; share no client; and have no formal technical education.

Three things are important to take away from this history: First, the technologies developed through the Free Software movement have routinely proven to be superior to, and more popular than, those developed by corporations and governments. Second, many participants in the Free Software movement have not gone through traditional engineering education, though some do so later. Third, Free Software has an ideological component, but it is also a grounded set of technologies and practices that have reduced the advantages that large-scale enterprises like companies or governments had in developing technologies.

On the first point, technical superiority, I was a close witness to this process: When I joined MIT in 2001, it was common for research projects to be built from proprietary systems like Microsoft Visual Studio (a programming environment), Access or DBase (databases), on closed operating systems (like Windows). In the last five years at MIT I have not seen a single project launched from proprietary software. F/LOSS software and open data has proven to be so technically superior that it has displaced commercial alternatives, and its influence is gradually moving outward from the tools of hacker production to higher level and more consumeroriented software, as evidenced by

Wikipedia, Firefox and many other crossover technologies.

The second point—that many Free Software participants have not gone through traditional engineering education—means that they have bypassed the inculcation described in the first section of this article. Not only have they not had to sit in seminars on how to deny their moral agency or made to choose regressive research projects, the Free Software movement offers new standards of exemplary engineering. Whereas emblematic programming languages might have once been the brainchildren of famous university professors (LISP at MIT) or industry researchers (C at Bell Labs), the new heroes are often independent or loosely institutionally affiliated, like Python's Guido van Rossum or Linux's Linus Torvalds, in high school when he began the Linux project. The distributed nature of free software has created an alternative structure of education and ethical inculcation_ to that of conservative engineering education. One can learn practical software engineering almost entirely online, through free books and tutorials or through intense social interactions on web sites like Stack Overflow or Git Hub, or IRC channels.

The third point is that the free software movement has developed a variety of concrete technology-augmented methods of collaboration. These include forms of self-governance, systems for managing many simultaneous authors (like version control or source management software), and even methods for resisting hostile opposition or sabotage (like the Debian initiation system). Many of these practices have been researched and described by management scholars. Baldwin (2000), for instance, has stressed the importance of design modularity that allows many simultaneous changes without worry about multiple changes conflicting, while Von Hippel (2005) has described how "user innovators" (like F/LOSS developers) are increasingly competitive with "producer innovators" (like companies) when communication costs decrease. Overall, these strategies and techniques are orthogonal to the means of technical production that defined the 20th century, namely the collocation of labor and capital in research universities, labs, and companies. Schools, labs, and firms are important actors in the free software movement, either as allies, hosts, or opponents, but ultimately free software works well without them.

We have seen how technology education and professional identity in engineering ultimately lead heterogeneous individuals to a dependent relationship with government and private enterprise, which in turn leads to the development of conservative technologies that reinforce the status quo. A new collective process of technology development, F/LOSS, offers an alternative to technology enculturation and thus liberates technologies from the goals of the most powerful in society. Technologies like Betweenthebars.org rely on F/LOSS not simply for the engines which make them run, but also for its model of productive, task- and product-oriented collective action and the accompanying techniques and software, like version control systems, that make F/LOSS possible. ☐

CHRIS CSIKSZENTMIHÁLYI *works at the intersection of technology and art and created the Center for Civic Media at* MIT.

WORKS CITED

Baldwin, Carliss Y., and Kim B. Clark. 2000. *Design Rules: The Power of Modularity.* Vol. 1. Cambridge, MA: Mit Press.

Kelty, Christopher M. 2008. *Two Bits: The Cultural Significance of Free Software*, Duke University Press.

Von Hippel, Eric. 2005. *Democratizing Innovation.* Cambridge, MA: MIT Press

ROMANS or BARBARIANS?

POLITICAL CAMPAIGNS AND SOCIAL MEDIA IN COLOMBIA

Elections are still about shaking hands and kissing babies, for the time being. Maria Vidart explores the first experience with social media campaigning in Colombia.

THE 2010 COLOMBIAN PRESIDENTIAL RACE revealed the tensions between not merely two candidates, but two different forms of leveraging political power. Juan Manuel Santos, the heir of two-term president Alvaro Uribe, ran against Antanas Mockus, the former mayor of Bogotá. Santos was portrayed as the continuation of the political pragmatism that was characteristic of the Uribe era. Santos' politics was associated with longstanding ideals of authority, security and order. In opposition, Mockus crafted an image of anti-corruption and transparency, and presented himself as an icon of political virtue. Accordingly, both campaigns had very different strategies to captivate voters. Santos' campaign relied heavily on local leaders and local politicians who mobilized their political structures to allocate votes, a process often typified and oversimplified[1] in Colombian public culture as clientelistic networks. Mockus' campaign appealed to a younger, urban, middle-class vote that was mobilized mostly via web-based social media.

By looking at how the two 2010 presidential campaigns in Colombia used social media, one can trace how notions about politics, individual freedom and political power are being negotiated in the era of rapid information exchange. Political managers in Colombia often make a distinction between "real politics" (disciplined, face-to-face networks of political allegiance where interests are negotiated), and "virtual politics" (political publicity and public opinion management). In their view, elections are only won through real politics, while virtual politics is an important aid, but one that does not win elections on its own. The use of social media for electoral purposes disrupts these neat definitions. Although social media are not political by definition—they can be used for many purposes—they allow something appealing for political and electoral mobilization: the easy and quick coordination of goals among large groups. In the 2010 Colombian presidential elections, the two main campaigns used social media very differently, therefore showing how social media are an open platform that enables different political models to become tangible forms of political practice. Santos' use of social media was highly managed and pyramidal. Mockus, by contrast, appealed to the spontaneous power of publicity and visibility that the unstructured use of social media generated. Networks such as Facebook and Twitter were awash with a collective enthusiasm for Antanas Mockus.

Closer to election day, Santos' team brought in political advisors in order to counteract Mockus's increasing popularity: the controversial J.J. Rendón as chief strategist and Ravi Singh, a self-alleged

---

1. An academic tradition, specially in Latin American political science, sees clientelism as a remnant of premodern and corrupt political cultures in underprivileged contexts. Clientelism is often understood as the exchange of interests, moneys and favors. Recent critiques argue for culturally nuanced understandings of clientelism as networks of solidarity and problem solving, a political practice that reproduces the State in a moment where state production is a complicated entanglement of private and public processes (Auyero 2000; Roudakova 2008; Schiller 2011).

(and fraudulent) social network guru.[2] Santos' strategy had to be redirected. Mockus was unstoppable. He had the support of young people upset with traditional politics and fighting for a radical change in the country's political culture. Mockus' way of doing politics reflected this sentiment. Supporters believed that arguments and sentiments representing a collective pursuit of the common good were best expressed in the openness of social media. Mockus' campaign viewed social media as a transparent space that anyone could oversee and join. In this space, the power of the crowd's will would mobilize a symbolic political repertoire in order to transform a traditionally corrupt form of political practice led by interest-oriented political brokers. In an innovative political bet, Mockus relied heavily on the publicity pieces that eager individual supporters designed and circulated on Facebook, YouTube, Twitter, and Flickr.

The reliance of Mockus' campaign on the spontaneous support from its followers in order to mobilize a political message of civility, transparency and legality expresses how social media are sometimes

---

2. Rendón was widely known in Latin American political circles and in political media as the master of negative publicity, a practice scorned in Latin American political culture. Ravi Singh misrepresented himself as a campaign guru who had won the Obama election, according to an investigative report that circulated months after the election (León 2011). Campaigns worldwide hired his services believing his credentials.

conceived for electoral purposes. The main appeal of this medium, according to Internet-campaigning pioneer Joe Trippi (2008), is the interactive interplay between message producers and receivers. The interplay establishes intimate connections between the people and the political message circulating. More than passive consumption, users actively define and engage with political campaigns, they actively participate in democracy, as Trippi enthusiastically puts it. This interpretation of social media for political purposes expresses a fervent belief in these media to be the ideal place for an individual exercise of freedom, for unmanaged, voluntary individual participation to become a collective free will capable of social and political renovation (as Daniel Kreiss also explores in this issue).

Nevertheless, in order to win elections in Colombia, open and voluntary web-based participation was not enough, and Santos' campaign knew it. On May 30th after the first round presidential election, the polling results showed that the enthusiasm for Mockus was a mirage. Santos won the election by a 25% margin. A public joke circulating the next day in the same social media that glorified Mockus sardonically described his defeat: "Mockus won in Putumayo [the one territory he actually won], Twitter and Facebook."

The day after the presidential election, I met with President San-

tos' main strategist, J.J Rendón. He celebrated victory with a glass of wine. "Mockus' followers were like a crowd of barbarians; he had many followers but they weren't organized. When I arrived to Santos' campaign I organized it, as Roman Generals organized their legions" Rendón explained. For him organizing the campaign entailed aligning political allies and closely observing the power networks of the country in order to work within them. As Rendón remembers, Mockus may have had the multitude and the publicity on his side, but Santos had politics on his.

Aligning the forces of key political figures with a strong electoral base was decisive for Santos' victory. As political managers often explain, social media users are not disciplined voters. They are voters that act upon emotion and their votes are highly unreliable. For instance, as Mockus' campaign showed, there is little control over such voters: most of the people who enthusiastically participated in web scenarios were underage and hence could not vote. In opposition, the votes of established electoral bases, those that are negotiated during 'real' in-person interactions and mobilized by recognizable political leaders, are stable and predictable, they are disciplined, like a Roman legion. Political managers code these differences as a problem of class. Social media users participate freely and therefore nothing binds them to their candidate. Electoral bases are

expected to participate out of need. Therefore, elections in Colombia are won through the kind of loyalty ties that form and the networks of favor exchange and negotiation of interest built over years between politicians and their constituencies.

The Santos campaign's use of social media followed this highly managed and vertical form of political practice. Alejandro, a member of Santos' social media team described his involvement, "when you spot a sentiment on the Internet, you only need to light a match and it starts spreading like a wild fire." The trick is to control the fire. He was in charge of lighting the fire and keeping it burning. Web developers replicated pyramidal campaign structures of administration that resorted to (human) gatekeepers of information in order to guarantee the efficacy of social media. The web administration team combined administrators and trustworthy members who spread information among their networks. Their task was to keep the campaign message "disciplined," in line with the rest of the campaign, and to regulate web-based participation to avoid any attacks from hackers.

In this sense, social media in Santos' campaign resembled public relations, "a vital tool of adjustment, interpretation and integration between individuals, groups and society," as Bernays (2004:7), the father of public relations writes. This view is similar to Mockus' understanding of social media in that it recognizes the importance of free and open exchange of information for political projects to survive. But, contrary to Mockus, this notion of publicity in politics stresses the need to manage the flow of information in order to keep a disciplined message that integrates individual with collective goals and emotions, as well as some sort of control in how public issues are perceived. Management keeps the information tight, mobilized strategically, as politics requires.

Electoral politics is intricately strategic (Bourdieu 1991), like a game of chess, and Santos' campaign knew how to play the political game. The campaign knew where political power lay and it wasn't on the Internet but within highly managed, tight, local political networks. However, the use of social media for electoral purposes in the presidential race of 2010 challenged straightforward assumptions about Colombian politics for anyone following the elections. For a brief moment in Colombian public culture, unmediated participation in social media opened the real possibility for anyone to jump in the chess game. Social media users saw the possibilities these media have for political renovation. Colombia saw how an amplified, collective voice could shake up established politics. Santos' campaign trembled with Mockus' social media popularity. But, after this event, savvy political managers, whose job is to navigate the intricacies of electoral politics, were quick to incorporate these new technologies to their own advantage, as another space where to reproduce established political logics. As events worldwide show, the barbarians await by the city's doors. The question is whether politics will really change when they come in. □

MARIA VIDART *is a PhD Candidate in Cultural Anthropology at Rice University.*

MOCKUS PRESIDENTE

"CON EDUCACIÓN TODO SE PUEDE"

Pemitor*

CITED WORKS

Auyero, Javier. 2000. *Poor People's Politics: Peronist Survival Networks and the Legacy of Evita.* Durham: Duke University Press.

Bourdieu, Pierre. 1991. *Language and Symbolic Power.* Cambridge: Harvard University Press.

Bernays, Edward. 2004. *Public Relations.* Whitefish: Kessinger.

León, Juanita. 2011. "El Mito Del 'hombre Del Turbante': Una Farsa." *La Silla Vacía,* April 24. http://www.lasillavacia.com/historia/el-mito-del-hombre-del-turbante-una-farsa-23481.

Roudakova, Natalia. 2008. "Media political Clientelism: Lessons from Anthropology." Media, Culture & Society 30(1): 41 -59.

Schiller, Naomi. 2011. "Catia Sees You: Community Television, Clientelism, and Participatory Statemaking in the Chávez Era." In *Venezuela's Bolivarian Democracy: Participation, Politics, and Culture Under Chávez* Pp. 104-130. Durham: Duke University Press.

Trippi, Joe. 2008. *The Revolution Will Not Be Televised: Democracy, the Internet, and the Overthrow of Everything.* New York: HarperCollins.

EVERYWHERE AND NOWHERE

FOCUS GROUPS AS ALL-PURPOSE DEVICES

Can a focus group be all of us? **Rebecca Lemov** explores how the box of donuts and the one-way mirror have become essential features of our self-understanding. ▸

By the end of the twentieth century, the focus group came very close to being a universal tool in America. Even in face of a generation of new attention-spangled, Twitter-spawned techniques, the focus group continues soldiering on as one of the most ubiquitous devices that affects and reflects on modern life in the early twenty-first century. Indeed, when implemented on such platforms as Facebook and Internet meetup sites, the focus group is continually subject to rebirth and reformation. It is uniquely malleable.[1] This under-the-radar quality allows it to adapt to new conditions in the so-called digital age. At least, this is what I mean to prove, but I will not be surprised, for reasons detailed below, if the reader does not immediately share this point of view, for focus groups, to many people, are a bit like mosquitoes: they certainly exist, and there is admittedly little we can do about that, but let us try not to pay them any more attention than they deserve, besides a distracted swatting.

Focus groups, however, are more complicated. Even their "nowhere-ness," their seeming inconsequentiality, is a key component of their power. One of the most interesting qualities of focus groups is that they seem so dull, conjuring up images of badly furnished rooms and fluorescent lighting. They seem, in short, hardly worth focusing on. Yet they are almost everywhere. (They are focusing on *you*.) How did this odd

now-I see-you-now-I-don't quality come about?

BORN IN WORLD WAR II as a device to gauge American GIs' attitudes toward the fascist enemy and the possibility of inculcating a greater willingness to fight, the focus group, then known as the "focused interview," was the contribution of sociologist Robert K. Merton. Earlier studies such as John Dollard's Army-sponsored and Rockefeller-funded "Fear and Courage in Battle" revealed that a surprising portion of United States troops could not articulate their reasons for going to war and this resulted in a less vigorous will to fight—in short, a less motivated soldier. How to remedy the troops' anagnosia—their ignorance of geopolitics and fascism in particular? Campaigns went forward on many levels including propaganda, education, conditioning, and—of concern in this current brief history—investigating the susceptibility of trainees' inner attitudes to persuasive message. First, an expert motivator had to see clearly "into the heads" of those he wished to address. As Merton recalled of his first encounters, in the 1930s, with the prelude to the focus group (interviews he conducted with Boston-area hobos and homeless men under "sometimes strenuous" conditions), "[T]his situation strikes me as providing almost privileged access to people's states of mind and affect" (Merton 1987: 553). He continued to develop

this interviewing method and, as the war went on, Merton used it to home in on soldiers' minute-by-minute responses to situations and information—to provide an ever-improving quality of data.

Along the way, Merton elaborated many new and targeted techniques. The focused-interview technique of "graphic reinstatement" produced the following exchange in wartime research:

Interviewer: **A little while ago, you were talking about the scenes of bombed-out school houses, and you seemed to have more ideas on that.**
[Show still from film.]
How did you feel when you saw that?
Recruit: I noticed a little girl lying under a culvert—it made me ready to go fight then. Because I have a daughter of my own, and I knew how I would feel if anything like that happened to her (Merton 1956: 60).

The technique allowed a researcher to pinpoint the instant at which and the exact visual cue through which a soldier's motivation to fight shifted. It was almost like finding an "on" button. But the point of the focused interview was not just to locate "on" but also, above all, to disentangle strands of motivation and nodes of concomitant behavior change. In other words, it was to be highly precise about the technique's operations.

---

1. Reports of its death, according to at least one blogger, have been exaggerated http://brandsavant.com/the-death-of-focus-groups/)

By the mid-1950s the research device, now named the focus group, found its way to the commercial sphere. (It was largely ignored in the nascent and munificently-funded Cold War behavioral sciences as "too qualitative.") Research from Harvard small-group sociologist Robert "Freed" Bales brought into circulation—in the form of his self-designated "Special Room" experiments—the now-familiar architecture of one-way mirror and bank of experts. A standard method was emerging, and according to my research, the combination of Merton's techniques with Bales' set up rendered a flexible design that could be constructed in almost any set of tandem rooms. The focus group's popularity spread. It was focus groups, or so legend as well as the notorious figure of motivational guru Ernst Dichter had it, that told marketers why instant cake mixes weren't selling to *Feminine-Mystique*-era homemakers. Such mixes, with freeze-dried egg product already included, required the baker to add only one thing: water. Yet American women, when they spoke up in focus groups, revealed they wanted both the advantage of a convenient mix and the feeling of having added something substantial. Marketers recommended, with a Freudian twist, that Betty Crocker alter its mix to require the addition of a fresh egg: "The industry recognized this feeling of guilt and said, 'All right, if you feel that bad about it, add your own eggs.' Now the housewife felt very happy because she could use the cake mix and still express her individuality," reported Dichter in his characteristically vague but decisive style (Dichter and Berger 2002: 157). In this way a bromide was born. Industry reasoning, supported by focus-group evidence, was that the housewife "needed permission" to indulge herself, which permission came in the form of the egg-based benediction—proof she was adding something of herself, not her ovaries but her "individuality." Accordingly, despite the tenuous Freud-slash-quasi-liberationist logic, sales of such cake mixes rose. By the time Sara Lee purveyed a line of instant confections in the 1970s, consensus had fallen on the winning formula of 70/30: 70 percent contribution from the cake-mix assemblage and 30 percent from the baker. The fact that this success story has achieved the status of apocrypha, although it seems genuinely to have occurred, is of interest.[2]

From testing kitchen products the device spread to the movie industry in the 1970s, and to PR and K Street in the 1980s and 1990s. Recently the focus group vaulted into the domain of "meta" or perhaps came to grips with its farcical tendencies when Stephen Colbert adopted focus-group guru Frank Luntz's techniques to brand his forthcoming SuperPAC.[3] A clip finds Colbert eating popcorn behind a one-way mirror while a roomful of potential voters debates the merits of his proposed slogan, "Corporations are People." This would seem to mark a fairly recent development in the focus group, in which they become the object of humor. In addition to inspiring attacks and high-level critiques, the focus group is now well known enough to be spoofed in an ad for fast food in which a participant talks back to the observers behind the one-way mirror. Focus groups' entertainment value is apparently self-evident: not only do clients watch actual groups in real time, but a 2010 "Turnaround" campaign for Domino's even included video footage that captures a focus group gone dramatically awry, with participants declaring, "Domino's pizza crust tastes to me like cardboard" and "Worst excuse for pizza I've ever had"—a story line shortly followed by on-camera Kamikaze visits to the erstwhile focus-groupers' homes to force them to taste the re-engineered pizzas.[4]

All in all, via the focus group participants are the objects of an intense *focus.* Each person, actually or metaphorically, is the target of a research team's experimental eye. The focus group, thus, is part of a process of exploring and colonizing more and more private realms, a process that now extends to the social bonds between individuals, themselves subject to scrutiny and what might be considered a relatively new form of control—one that approaches "ultrarapid forms of free-floating control" (Deleuze 1992: 3). As technical means of "getting in tune with the reality of the interviewee" (Kreuger and Casey 2009: 3), focus groups share the central goal of the social sciences—to see the world as if from another person's point of view—while at the same time they constantly calibrate a new way to encounter that world. They change what they know. They intervene in reality.

EVERYONE KNOWS WHAT "focus group" means, or at least thinks they do — it is at root a specially designed room fully rigged with microphones and a conference table, possibly equipped with a box of donuts, and featuring one-way mirrors instead of windows. Behind the mirrors sit the expert researchers, tracking participants' behavior, taking copious notes, and deriving conclusions. A gathering of clients or business owners or campaign operatives sits in an adjunct space watching as, inside the room on the other side of the mirror, a group of people assemble around a table tasting salad dressing, speak-

---

2. Still, whether it occurred at Betty Crocker or at Bisquick is not clear. See, e.g., the account in *Newsweek:* "Inside the Consumer: The New Debate: Does He Know his Own Mind?" *Newsweek* 10 October 1955, p. 92, which centers on biscuits.

3. http://www.colbertnation.com/the-colbert-report-videos/394777/august-16-2011/colbert-super-pac---frank-luntz-commits-to-the-pac

4. http://www.youtube.com/watch?v=AH5R56jILag

ing about a candidate, trying out a cell phone, responding to a public policy, or watching a film. As a result, a somewhat randomly gathered quorum of people, preferably though not necessarily strangers and distinguished in part by their willingness to attend a focus group—their views duly recorded, transcribed, and interpreted—can thus have an inordinate effect on the national direction. Political steering by means of focus groups inspires derision from some quarters, as when Joan Didion, speaking of how focus groups could be used in 2004 to prove one thing or, equally, its opposite, commented: "The reduction of the American electorate to twenty people who lived in or near Cincinnati was in fact the elegance of the mechanism, the demonstration that the system was legible, the perfected codex of the entire political process" (Didion 2004). It is not only an act of reduction that focus groups accomplish—from two million potential voters to twenty Midwesterners as their synecdochal voicebox. As Didion pinpoints, the focus group's work is also to *stand for:* it embodies the "perfected codex" of the political system itself. It actively symbolizes that which it represents. It marks and sustains a kind of confidence and streamlines the dimension of trust in which, in fact, it traffics. For with its circular mechanism, the technique renders reliable that which it produces, even when it is wrong.

Democratic devices, they inspire more and more talk from the subjects on whom their focus trains. *Thank you for listening to me,* researchers report hearing when a group has finished its session. "There is something about a focus group that is different and causes people to *feel* that someone is listening," comment two veterans (Kreuger and Casey 2009: xiii).

THE PROCESS OF RUNNING focus groups and making decisions from their operations helps reinforce their own conditions of knowledge, reminding me of what Pierre Bourdieu said about habitus, or the conditioning conditions that make up social life itself: they function like a train laying its own tracks. In a paradoxical manner, the focus group runs along tracks its own movement lays out, as it were, post-hoc but somehow preemptively. Another way of saying this is that a focus group is a little engine for targeted change. A focus group can make a presumptive case, preemptively-arrived-at, into an unavoidably-so reality. They not only change policies and press relations, they change the people who participate in them. People come out different from how they went in. The focus group itself has become an agent of causation, a scripter of response. It shows eighteen-year-old sweat-shirted boys and married mothers-of-one *how to talk* about products, *how to relate* to things or phenomena, and are thus social engineering steering devices, in effect. As researchers Bristol and Fern have shown, participants in a focus group will experience a unique mix of "anonymity and arousal"— which, better than any other method, facilitates "expression of shared experiences." (Bristol and Fern 1996). Focus groups address the realm of the scarcely thought out, the inchoate, the shared but not spoken social reality between and among people–and then they *alter* these hard-to-talk-about things. Focus groups thus give a feeling of inevitability to the previously contingent. They are not just about representing a truth about reality. They change reality.

To return to where we began, there is a curious paradox inherent in focus groups. On the one hand, they are not a secret: While conducting this research, I was often told, "Oh, focus groups—my sister runs them"—or a cousin hires them, or a student friend does mock juries for them, an aunt makes her living participating in them. Yet for all their ubiquity they are strangely unknown. Their demise and obsolescence in an "Internet age" is frequently announced, yet their popularity with corporations and campaigns remains unabated, with or without online techniques tacked on. They are at work pretty much everywhere, but few know *how* they work (even those who administer them) and even fewer know the unlikely history of where they come from. They constitute a vast social experiment the results of which are unknown. ☐

REBECCA LEMOV *teaches in the History of Science Department at Harvard University.*

WORKS CITED

Bristol, Terry and Edward Fern. 1996. "Exploring the Atmosphere Created in Focus Group Interviews: Comparing Consumers' Attitudes Across Qualitative Techniques," *Journal of the Market Research Society,* 38: 185-195.

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

Dichter, Ernst and Arthur Berger. 2002. *The Strategy of Desire* Transaction Press.

Didion, Joan. 2004. "Politics in the 'New Normal' America," *New York Review of Books,* Oct. 21; http://www.nybooks.com.ezp-prod1.hul.harvard.edu/articles/archives/2004/oct/21/politics-in-the-new-normal-america/

Krueger, Richard A. and Mary Anne Casey. 2009. *Focus Groups: A Practical Guide for Applied Research.* Thousand Oaks: Sage.

Merton, Robert K. 1987. "Focus Groups: Continuities and Discontinuities" *Public Opinion Quarterly,* 51: 550-566.

Algorithmic Recommendations and Synaptic Functions

Personalized recommendation is the new marketing. Nick Seaver explains how 'collaborative filtering' defines people through their purchases.

In her ethnography of a Norwegian marketing firm, the anthropologist Marianne Lien describes an advertising campaign intended to promote frozen pizzas:

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

Pizza Superiora was "the people's pizza," intended for a general audience; Pizza Romano, more expensive and with "a distinctive flavor and character," targeted "a more adult and selective audience"; Pizza Preciosa, with a wholemeal crust and vegetable topping, was aimed at "women aged 15-40 focusing on health, body and appearance. Vegetarians" (171-2). The work of the marketers, Lien argues, was to forge connections between their products and groups of consumers. A successful connection between a market segment and a frozen pizza product would lead to economic success.

However, as Lien shows, people and products are not stable entities. Market segments and frozen pizzas change over time and in response to each other. As marketers tailor product lines, consumers buy and eat with an eye to their own social distinction. Successful marketing campaigns do not only identify "real and potential target groups" — they produce them. The Pizza Romano and the "more adult and selective" pizza eater emerge in concert with each other. Intersections of gender, age, and dietary restriction are made meaningful by the differentiation of targeted products. In demographic marketing, groups of people and groups of products are mutually defining: brand strategists understand pizzas in terms of people and people in terms of pizzas.

SYNAPTIC FUNCTIONS

In his 1978 book *Culture and Practical Reason,* Marshall Sahlins draws a provocative comparison between

the work of marketers and the work of social scientists:

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

According to Sahlins, marketing is itself a kind of social theory — a mode of sensitivity to "latent correspondences in the cultural order," organized around the imperatives of commerce. Evident in Lien's account of the Norwegian firm, demographic marketing is a way of understanding groups of people through their correspondences with groups of things. And, although they do not produce symbols "de novo," it seems clear that marketers not only describe latent correspondences — they, at least in part, create new ones as new products come to market and become vehicles for the expression of social distinction. The synaptic function is both descriptive and generative.

In order to examine the generative qualities of the synaptic function, I describe here a contemporary challenger to traditional demographic marketing: an algorithmic recommendation technique called "collaborative filtering." Collaborative filters are an increasingly frequent feature of online infrastructure, suggesting books, movies, music, and news to users. They draw correspondences between users and items by comparing user ratings, producing recommendations with the familiar form: "Users like you liked items like this." In performing their synaptic function, social theorists (such as marketers and anthropologists) endorse and generate new figurations of social collectives. Novel modes of understanding correspondences between persons and things, it follows, may produce novel figurations of social form.

"ON THE INTERNET, THERE'S NO EXCUSE FOR NOT PERSONALIZING"

In 2002 John Riedl and Joseph Konstan, a pair of computer scientists from the University of Minnesota, published a book for marketing executives. *Word of Mouse: The Marketing Power of Collaborative Filtering* promised to upend the marketing world by sharing the secrets of a new science for understanding consumers — not as members of demographic groups, but as individuals. "The urge to poll and classify is intoxicating" (109), they wrote. "The problem is, simple demographics don't begin to tell the story of individuals" (112). With the advent of online retail and new technologies for tracking the activity of customers, marketers could begin to follow these individual stories, targeting users not through generic demographic profiles, but

Word of Mouse: The Marketing Power of Collaborative Filtering

with personalized recommendations.

Collaborative filtering was an algorithmic technique for producing such recommendations. Users would rate items (explicitly or implicitly, e.g. on a 5-point scale or by purchasing a particular item), and on the basis of these ratings, the filter would make suggestions drawn from the ratings of similar users. These similar users, algorithmically determined, took the place of market segments. Instead of assuming that a customer would want what others in their demographic group wanted, the collaborative filter assumed that customers who shared some preferences would also share others. Part of collaborative filtering's appeal was its economy: the only information it needed to work was a set of numerical ratings — information about specific users or items to be recommended was superfluous.

The book's cover promoted collaborative filtering as the equivalent of ESP for Sahlins's synaptic function: "Know what your customers want even before *they* do." Below that slogan, a cheery and diverse crowd of customers waved from

▶ **SPARSE:** An example of the archetypal Collaborative Filtering Matrix.

USERS →	A	B	C	D
ITEMS ↓				
(adidas shoe)	★★★★	★★★☆		
(red Nike shoe)		★★★★		★☆☆☆
(dark Nike shoe)			★★★☆	★★★★
(blue Nike shoe)	★★☆☆	★★★★		

inside a computer monitor, apparently pleased by this technological breakthrough in taste prediction. This group represented the collaborators of collaborative filtering — the users whose aggregated activity could be algorithmically mined to predict each other's preferences.

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

users do not know or communicate directly with one another, through the algorithm they are made collaborators — a computationally arranged aggregate of taste-bearing individuals.

MAKING SIMILARITIES IN THE MATRIX

In order to understand the kinds of groups made and understood through collaborative filtering, it is essential to wade into its technical form — the algorithms tasked with finding order among individuals.

The archetypal form of a collaborative filtering system is a matrix: a grid, with items along one side, users along the other, and ratings at their intersections. This matrix is mostly empty (or "sparse"), since most users will have not rated most items. The work of the collaborative filtering algorithm, as it typically stated, is to predict what values will show up in the empty spaces of the matrix. These predictions are then provided in some form to the user as recommendations. Thus, at any given time, the matrix is in an anticipatory flux: new ratings from users arrive constantly, displacing their predicted values and shifting the others. This filling process is the signature action within the matrix

— blank values are replaced by predictions, which are then replaced by actual ratings. Progress from emptiness, through prediction, to actualization makes the matrix a proleptic social representation, holding simultaneously a record of past correspondences between persons and things and the anticipation of future ones.

The collaborative filtering matrix intermeshes the identities of users and items. It is both possible and typical for a collaborative filter to take no special account of either, organizing all entities strictly in terms of ratings: users are known as a collection of relations to items and items are known as a collection of relations to users. Persons and things enjoy no separate modes of existence in the matrix, which is indeed a function for translating one into the other: consumers can use the filter to organize items, and marketers can use the filter to organize consumers.

A common approach to recommendation is illustrated in this diagram from a recent article: the numbers from the matrix are statistically analyzed and their variance is mapped to a number of axes (in this simplified illustration, only two). Users who are near each other on this coordinate system are similar, and a user will be recommended items from the "neighborhood" around them. Although the axes that represent latent factors need not be labeled in order to produce recommendations, labels are often used as a way to explain and justify a system's output. Here, one can see the persistence of demographic ways of understanding groups: this figure organizes its contents according to gender and seriousness, making sense of the algorithm's output through its similarity with conventional ways of categorizing movies. The diagram also makes evident the role of the word "like" in "Users like you liked items like this": preference and similarity are collapsed in this coordinate system, where "being like" and "liking" have been equated. You
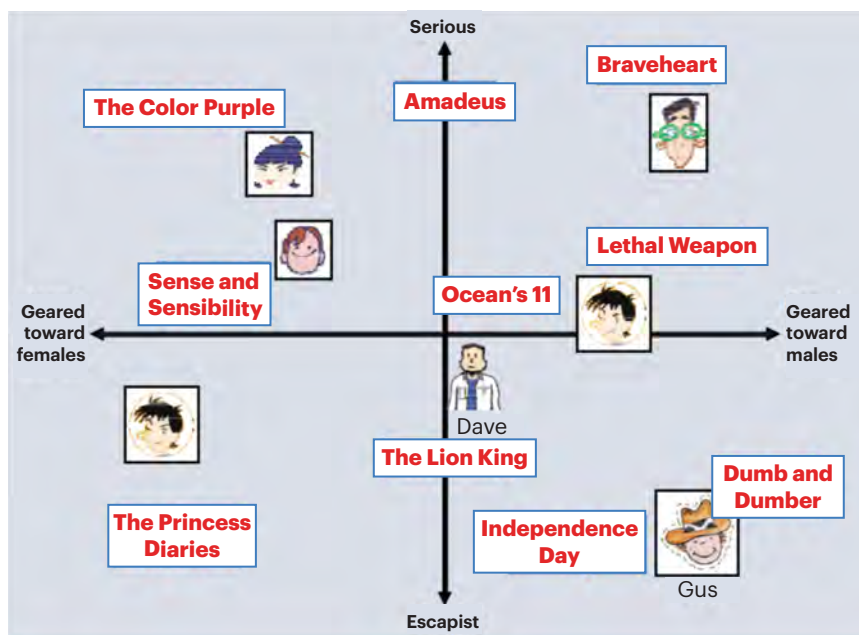

Figure 2. A simplified illustration of the latent factors approach, which characterizes both users and movies using two axes—male versus female and serious versus escapist.

may not like the same things as the rest of your demographic group, but you probably will share preferences with your "nearest neighbors" in the abstract cartography of collaborative filtering.

AUTOMATIC CORRESPONDENCE

The contrast in *Word of Mouse* between individualized recommendation and the "lazy, prejudiced philosophy" (113) of demographic profiling was in fact a novel frame for collaborative filtering. Publications from Riedl and Konstan's research group at the University of Minnesota and other groups around the US working on recommender systems had tended to emphasize the "collaborative" in collaborative filtering. These systems were envisioned as a way to reconnect lone users to larger groups, to "automate word of mouth," as one paper put it (Shardanand and Maes 1995), and to mediate between individuals and the unwieldiness of increasingly large data sets. Where *Word of Mouse* championed the re-emergence of individuals from the blunt taxonomy of traditional marketing, the academic literature on recommendation often focused on the rearrangement of those users into more meaningful groups. Collaborative filtering was not about privileging individuals over broader demographic categories, but about reinstalling isolated individuals into an algorithmically tuned collective.

In the techniques that collaborative filters use to organize individuals into collective forms, Sahlins's comparison of social scientists and marketers has taken a more literal existence. The statistical methods that populate the matrices of recommender systems, such as correspondence analysis and its variants, are commonly adapted from the social sciences. The famous graphs of Bourdieu's *Distinction,* coordinate systems depicting the correlations of taste and class, derive from methods remarkably similar to those that now power many online recommenders (for more on the history of correspondence analysis, see Desrosières, this issue). Collaborative filtering automates the synaptic function, moving the correspondence-finding work of social theorists (such as marketers and sociologists) into the algorithm.

This automation of social theory has a number of potential effects. As algorithms, specific theories about the correspondences between persons and things can be built in to the infrastructures of browsing and purchasing online. Used as filters, social theories become increasingly performative: the models of social science may come to shape the phenomena they were meant to describe. Through the continuous collection of user data, the collaborative filter has an increased level of flexibility and responsiveness: the positions of individuals vis-à-vis groups can change continuously. John Cheney-Lippold has recently described this kind of algorithmic interpellation as "soft biopolitics" (2011), a shifting mode of categorization that necessitates a reappraisal of the models of power and taxonomy in Foucauldian biopolitics. As a result of their flexibility and infrastructural existence, there is a risk that these systems will evade critique, coming to appear as natural and objective methods of organization.

Algorithmic recommendation is not simply a higher-resolution representation of a market — a more precise picture of atomistic individuals that does away with the need for larger-scale approximations like market segments. Rather, it is another mode of the synaptic function — another technique for making and interpreting correspondences between persons and things, another way of organizing collective forms. Collaborative filters algorithmically rearticulate the relationship between individual and aggregate traits, suggesting the need for social scientific theories that eschew the classic break between groups and their members (for a preliminary attempt at such an approach, see Latour et al., forthcoming).

The work of recommendation, like the work of demographic marketing, relies on the idea that there are meaningful similarities among consumers and that these similarities correspond with similarities in objects. However, in algorithmic form, these correspondences take on new forms and meanings, blending preference, identity, and similarity. As these theories are built into online infrastructures, shaping the relations between persons and things and articulating new collective forms, they demand attention, not only as material for analysis, but as new modes of analysis itself. ☐

NICK SEAVER is a PhD student in the Department of Anthropology at the University of California, Irvine.

WORKS CITED

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste.* Trans. Richard Nice. Cambridge, MA: Harvard University Press.

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

Koren, Yehuda, Robert Bell and Chris Volinsky. 2009. "Matrix Factorization Techniques for Recommender Systems." *Computer* 42(8).

Latour, Bruno, Pablo Jensen, Tommaso Venturini, Sébastian Grauwin, and Dominique Boullier. n.d. "The Whole is Always Smaller Than Its Parts: A Digital Test of Gabriel Tarde's Monads." *British Journal of Sociology.* http://www.bruno-latour.fr/sites/default/files/123-WHOLE-PART-FINAL.pdf

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

Riedl, John and Joseph Konstan. 2002. *Word of Mouse: The Marketing Power of Collaborative Filtering.* New York: Warner Books.

Sahlins, Marshall. 1976. *Culture and Practical Reason.* Chicago: University of Chicago Press.

Shardanand, Upendra and Pattie Maes. 1995. "Social Information Filtering: Algorithms for Automating 'Word of Mouth.'" *Proceedings of the SIGCHI conference on human factors in computing systems.*

PUBLIC SAFETY & WALL STREET

Compstat and the **Real Time Crime Center** are at the epicenter of Bloomberg's New York. *Emmanuel Didier* explores how they are turning public safety into a commodity for Wall Street.

IN 1994, RUDOLPH GIULIANI was elected Mayor of New York following a campaign centered entirely on the problem of public safety. He promised New Yorkers that he would restore the "quality of life" which had been destroyed by criminality. This promise was certainly directed at the lower and middle-classes, who were suffering so badly that they started to leave Manhattan (in preparing for the campaign, Giuliani ordered studies on the reasons behind this sort of migration, both inside and outside New York, that showed just this). More implicitly, this promise was directed at major international business players in the "big FIRE" sector ("Finance, Insurance, Real Estate"). True, this group had the means to protect itself. They lived and worked in ultra-secure areas. They were not personally threatened. But their status as the global elite gave them a special quality of life comprised of excellent restaurants, high-end clothing stores, elegant French bakeries, etc. Now, these services are rendered by none other than those very lower and middle classes who were directly suffering from crime. If the middle class left, the powerful elite of New York might do so, and risk that Wall Street would pack up

and move to London or Tokyo. The city would no longer be the "sustainable" Mecca of world finance. The fight against street crime was therefore also indirectly about permitting New York to continue to amass immense capital. As surprising as it is, the argument is not mine, but that of a number of actors in and analysts of New York (Sassen 2007, Vitale 2008, Interview with O'Boyle 2010).

After the September 11 attacks, all the attention was shifted to the fight against terrorism, which completely overshadowed the link between Wall Street and the regulation of public safety in New York. Nevertheless, this regulation has undergone two successive and important revolutions revolving around the very statistical and information technologies it depends on. Almost immediately after being elected, Rudolph Giuliani recruited William Bratton, who put into place a technology called Compstat which has been widely talked about – even among the public at large since it was featured in the television series *The Wire,* set in Baltimore, which like other cities the world over had starting using it. The Compstat revolution reconfigured crime into *rates of variation*

from statistical indicators (basically crime statistics), and the acting police chief was responsible for decreasing or increasing these numbers—and proving just that before the city council in public meetings.

Then, in 2005, new mayor Michael Bloomberg willingly accepted police chief Ray Kelly's implementation of another device called the Real Time Crime Center (RTCC). The tool is not strictly statistical, in that rates and quantities are replaced by giant, widely varied *databases* with which detectives do *data mining* and engage in real-time identification of individuals who have just committed crimes.

Despite these temporal coincidences, the connection between finance in New York and the new public security measures is rarely noted. Having focused for several years on the implementation of Compsat in Paris (Didier 2011a, 2011b), I went to New York in April, 2010 to observe the original version of this technology, which we have adopted in France and there conducted this ethnography on police matters. I met numerous NYPD officers of all ranks, visited a police precinct headquarters in Manhattan, and talked with the upper echelon in the force, including Stephen

Goldsmith, then deputy mayor, who later allowed me to visit the RTCC in NYPD headquarters at the legendary 1 Police Plaza. I argue here that these two technologies complement one another by reconfiguring the connections between public safety policing and the traditional collectives it is generally associated with: suspects, policemen, and the masses. And, at the same time, to try to support the unexpected hypothesis that these transformations are also connected to Wall Street.

COMPSTAT

Right after his election, Rudolph Giuliani recruited William Bratton, whom he had met when Bratton was Chief of Police in Boston and later Chief of the NYC Transit Authority. As police commissioner, he was in charge of reducing crime both quickly and forcefully. Bratton's response was to invent Compstat (well analyzed by Silverman, 1999), which he said came from two principal sources of inspiration. The first, which was extremely well received in business circles, was Hammer and Champy's *Reengineering the Corporation* (1993), a management theory from the private sector that defended the idea of placing individual responsibility with the officers and shortening hierarchies whenever possible. The other was the "broken window" theory of James Q. Wilson and George L. Kelling (1982), members of the very conservative Manhattan Institute. It stated that police must respond to all urban disorder, as trivial as a broken window, or else inhabitants would get the impression that their neighborhoods were neglected and would in turn neglect it themselves, abandon it, and finally leave it to real criminals.

In concrete terms, the precinct commanders are at the heart of Compstat. Halfway up the chain between beat patrolmen and police headquarter masterminds, these commanders were given as much power as possible to help implement their own crime reducing strategies. Thus, they had to show initiative. But all this freedom came with serious responsibility: every two months they had to go to a "Compstat meeting" to give an ac-

counting before a mixed assembly sometimes including the mayor himself, always the head police chief, other precinct commanders, and finally local residents and local business representatives.

The report basically meant preparing number charts and quantitative indicators for projection onto a giant screen. This stage was difficult because initially the data for each of the precincts was simply not available. The first step was therefore to become computerized (thus the name COMPuterized STATistics). Once this was accomplished, precinct commanders could use information much more quickly to show the public that police activity had increased (the number of patrols, time spent on the street, "stop-and-frisks" etc.) while crime had abated (whether for small infractions – for example, the campaign began with a war on squeegee men in front of Grand Central Station – or any of the "7 major felonies").

One must understand that criminal investigation divisions, which fight big crime, international networks, and the mafia, have never produced any important numbers because the enemy, although powerful, is a very small crowd. Arresting Al Capone is striking only one arrest. On the contrary, the broken window theory, armed with Compstat, allows for one to compile, add, and present very important misdemeanor numbers.

Now, Compstat serves not only to represent crime, but also – and especially – to act upon it, transform it, and reduce it. What's more, Compstat was often considered a miracle because these misdemeanors, which initially appeared in staggering quantity, truly began to subside. Many who lived in New York in the 1990s will tell you today that they feel much safer. This is what the police have to show: that small-time crime exists, that they actively fight it, and that in stemming its tide they also help to reduce violent crime. The collective police force is thus also reformatted towards an increased *subjectification* in the sense that they must show that they are taking initiatives, that they are inventive and proactive, and at the same time responsible for both suc-

cesses and failures. And responsibility is not an empty word because a good fraction of the commanders were disciplined after the first year.

Finally, these numbers are also clearly associated with a spectacle that makes the public take notice of municipal action. Compstat is not a system of dissimulation for the police, but just the opposite: it's an echo chamber. The neighborhood in question is invited to personally attend the precinct commander's presentations (though this is not true anymore, as they felt that humiliation was not bearable when they were being shouted at). The press is also invited to attend whenever it desires. Police officials from around the world crowd together and attend strategy meetings (in regards to France, consult de Maillard, Le Golf, 2009, which contains an impressive list of political representatives who have come to experience Compstat firsthand). Finally, the fact that Compstat appears several times as a setting in the TV series *The Wire* further illustrates its popularity. It's not enough for crime to subside; it's also necessary to furnish citizens with proof that helps convince and reassure them.

On might recall that scene from John Ford's *The Man Who Shot Liberty Valence,* when a journalist declares about the eponymous crime: "This is the west, sir. When the legend becomes fact, print the legend." Today we might say: "This is New York, sir. There is no longer any difference between fact and legend." The numbers are compiled in order to meet three objectives at once: first, to produce facts; second, to act upon and transform these facts; third, to create a spectacle around them. It's not about accusing the NYPD of lying, because in spite of the criticism, which has for the most part attributed the lower crime rates to factors other than police innovation, the crime rate actually has dropped. It's just that the tool used to accomplish this also allows for the creation of a legend that pleases everyone, including middle-class New Yorkers, the global elite, and police administrators the world over (Salmon, 2007).

Still, for a while now Compstat has been having problems (Eterno

and Silverman, 2012). The main trouble is that the need for continually lower numbers means that these very numbers have become less meaty. On the one hand, not doing better than in the previous year means admitting failure. One the other hand, truly doing better seems to have become more and more difficult - more difficult, in any case, than... cheating with the math. The Schoolcraft scandal is a reminder of this. It concerns an NYPD commander who grew weary of orders that led to an erosion of real police work at the price of obtaining better statistics (engaging in unfounded arrests and unjustified identity surveillance, refusing to take complaints). The commander therefore decided to record all his activity including unwarranted orders from superiors for several months, and after failing to get the attention of the upper echelon of the police department, he decided to give the recordings to *The Village Voice,* a New York publication with huge distribution, which published a series of very noteworthy articles, and became later the subject of an episode of *This American Life.* He is currently at odds with the NYPD in a court trial. One might wonder if this wasn't the beginning of the end for all this technology. But Bloomberg's team seems to have anticipated these difficulties, for in 2005 it developed a new police technology.

TRADE FLOOR MAYOR
The open floor offices of New York Mayor Michael Bloomberg was modeled on the trading floor. PHOTO BY THE ARTICLE AUTHOR

REAL TIME CRIME CENTER
Elected mayor in 2002, Michael Bloomberg, founder of the finance information network that bears his name (and also, incidentally, the city's wealthiest taxpayer), can not be accused of having left Wall Street in the lurch. He in no way questioned the link established by Giuliani between public safety and finance. On the contrary, he used the latter as a model for his own administration and physically transformed the mayor's office into an *open floor* based on the trading floor. He also did not question the use of quantification of officer activity as a means of evaluation. In fact, he expanded this technique for use in every department and sub-department in his administration.[1] Finally, as a journalism expert, he did not question the bombast with which the statistics were published. Still, he and Police Commissioner Raymond Kelly wanted to implement their own anticrime apparatus. They didn't really abolish Compstat, which was still operational after some minor modifications. Rather, the two technologies evolved together.

Bloomberg and Kelly, for their part, did not rely on a theory of management, as Bratton did. They went to an expert, Jim Onalfo, who had over thirty years of experience in enterprises ranking among the top 100 in *Fortune IT.* It was he who was charged with the task, and in 2005 the Real Time Crime Center (RTCC) was inaugurated.

Physically speaking, the center is a large single room at the police headquarters on 1 Police Plaza where work stations, all in a row, sit facing a giant screen ("data wall"). The 45 "investigators" and 16 "supervisors" working there were recruited after an internal job search at NYPD. They were hand-picked and presented as an elite team, and then split up in the room according to a hierarchy: detectives up front, then sergeants and finally lieutenants. They all serve as back-up for the field detectives. When the RTCC was created, these detectives were obliged to use make use of it, to get familiar with the system. Now it is apparently at their beck and call, but there's a waiting list.

The power of the RTCC lies in the fact that all these work stations are connected to a giant data warehouse (which I was not permitted to see). The data comes first of all from the police force itself. One of the innovations at the heart of the cen-

ter is that since Giuliani's tenure, the data for all New York precincts has been interconnected. Thus, requests can be made in all 911 calls, complaints, arrests, and narcotics files. 23 supplementary system files, some from New York's city and state records, are also interconnected: 311 (government directory assistance), Stars (parking tickets), e-justice, parole and probation, etc. Plus, the RTCC purchases public databases, like Lexis/Nexis, and also uses Google Earth. Finally, access to the files of telephone operators has been made possible without even requiring a subpoena from a judge. In the civil framework, this file exchange is comprised of "memoranda of understanding" between contacted parties. Facial recognition from video may still not be entirely possible, but the quantity of data available since implementing the RTCC is simply staggering.

The RTCC uses this information not for minor misdemeanors, which was Compstat's chief concern, nor against white collar crime, which is handled by another sector. It tackles mainly what are called the "7 majors": (homicide, rape, robbery, assault, burglary, grand larceny, arson). When a crime is reported, the detective on the scene can call the RTCC, which starts an investigation as quickly as possible in order to connect the observable elements at the crime scene to all pertinent information on file. This process of "connecting the dots" helps identify the suspect. A program called Cognos allows for queries into all the databases. For example, one can do a "pattern analysis" and connect cases that are similar in terms of modus operandi. One may also do "entity analysis," which is to say looking at cases that are similar while not yet identical. For example, one might compare the case files of "Jim Beam, born 3/12/1945" to that of "Tim Beat, born 2/12/1945" and that of "Joachim Beam, born 3/12/1944." All this data is sent back to the RTCC and projected onto the screen so that everyone working the case can take advantage of it. IBM, which furnishes the majority of the programs and hardware for the center highlights the fact that its tools can now handle the

semantics of data (the meaning) and not simply the syntax (the organization). According to Captain Godek, who showed me the RTCC, it's nothing less than the reproduction of human thought, only faster and made up of a mutual pool of investigators.

Here, the databases that Compstat relies on are also used, but they don't serve as support for a detailed breakdown. The objective is not to look at how many times an identical event has happened, but to bring together different and unique characteristics that will lead to identifying a single individual. Data mining does not count lists of identical modalities, but rather looks to cross different variables to identify a point of intersection. The list of suspects, therefore, is phenomenal in size, so exhaustive even, one might fear or hope (depending on whether or not one trusts the police) that escape is possible for no one. Compstat also makes an inventory of suspects for relatively important crimes, but it is based on a list of those already arrested for minor infractions. The RTCC enlarges the perimeter of suspicion to include all those who figure on the very long (nearly exhaustive) lists on these databases.

The police officers are now much less individualized than when they were present at the Compstat "Crime Strategy meeting." The RTCC on the contrary functions to pool

their thoughts. But the workers with access to RTCC constitute an elite force of limited number, making the process of subjectivation of the agents clearly less necessary. The selection of a qualified and elite group is also, such as it is, another sort of delegation of responsibility. What's more, it's the detective on the beat who gets credit for the collar if all is successful, and not the RTCC inspectors, who have no need for this. One consequence of this is that there are no statistics that illustrate the effectiveness of the center itself. We don't know *how* useful it has been. Still, all the presentations of the RTCC highlight cases of arrests made possible thanks to the center (d'Amico, 2006). The proof of its effectiveness lies in the accumulation of individual cases, and not in the addition of figures.

Finally, the RTCC most assuredly does not forget to put it all on parade. The conference hall in the Center, which even a French sociologist can enter, is also obviously open to the press, who can thus witness live the resolution of a particularly spectacular event. And herein lies the other virtue of the "case study":

REAL TIME CRIME CENTER
In a large single room at the police headquarters on 1 Police Plaza sits 61 detectives, sergeants, and lieutenants behind their work stations, all in a row, facing the "data wall". Real time and archived statistics are piped in to this room through a giant data warehouse, and analyzed by people and computers. The data is then used to help police officers in the field to identify suspects. PHOTO BY THE ARTICLE AUTHOR

it is much more usable for the press than some rate of statistical variation. It can easily be transformed into a short film to be shown on the internet or television networks, which are themselves simultaneously watched by those working at the RTCC, a fact attested to by the large number of TV screens found there. Furthermore, the cases are presented in the same language that the global elite learn in business schools: their education is also filled with "case studies" and they reason this way. Also, in our opinion these cases are the yardstick with which "Real Time" is measured, for after all, this "real time" refers just as much to the time involved in the crime and it's resolution, the time it takes for the press to give an accounting of it, and beyond that, the time elected officials require for communication. Here, legend and fact overlap; New York is not only safe, but is known for being safe.

CONCLUSION

For the last twenty years, public safety has been an important political theme in many of the capitals of global commerce. Most of the time, it is portrayed as good and right in and of itself, as if its principle function is to respond to the anguish of the middle-class citizens directly touched by crime. Safety is a public good because voters complain about crime.

But this sort of obvious justification ignores the fact that there may also be another public that is just as important, one that while not directly affected by criminality and not necessarily named, is nevertheless also the target of this discourse. In the case of New York, the global elite behind the "big FIRE" might well represent these silent interlocutors. It has not been possible to penetrate the arenas where these power elite – politics, police and finance – actually meet but the characteristics of the security technologies implemented in NYC confirms the hypothesis.

Indeed, over the last twenty years, the administration of public safety depends on two information technologies: Compstat and the RTCC. They are certainly different, but both can be connected to neo-liberalism as defined by Foucault (2008, 2009). In effect, we have seen that they function first of all to define a suspect "population" that police can identify and control. Furthermore, the police officers become "entrepreneurs of the self," which is a paradox for agents whose actions are supposed to be subject to the law. To this can be added a third policing characteristic not noted by Foucault but which still plays a major role in the mechanism: the ritual through which the numbers are presented to the public. The neo-liberalism that governs the police and that which the financial world has dedicated itself to are one in the same. And all the work done by the police administration, all the consequences thereof, lead straight to the ears of the big shots that comprise "FIRE." This is something that they must be satisfied with so that the rhythm of Wall Street remains stable, and the main players stay in Manhattan.

The problem is that when we create a public good for a certain group without saying so, it's likely that this group benefits from said good without paying for it. The accumulation of capital made possible through the public security of the NYPD is thus countered by the fact that the other beneficiaries of safety, the much less wealthy, are the ones who pay the price (and School-craft is there to remind us that the price is not only monetary, but also represented by an abuse of power, police harassment, and pressure). Safety is a public good. New Yorkers are offering it up to the new global elite in hope that they deign to stay and redistribute the benefits. ☐

EMMANUEL DIDIER *is a permanent researcher at EHESS - CNRS, Groupe de sociologie politique et morale (GSPM), in Paris.*

Translated by Paul Knobloch

WORKS CITED

D'Amico, J. 2006. "Stopping Crime in Real Time," *The Police Chief,* 73(9), September. http://www.policechiefmagazine.org/magazine/index.cfm?fuseaction=display_arch&article_id=995&issue_id=92006

Didier, E. 2011a. "L'Etat néo-libéral ment-il ? "Chanstique" et statistiques de police," *Terrain,* 57:66-81.

Didier, E. 2011b. "'Compstat' à Paris : initiative et mise en responsabilité policière." *Champ pénal/Penal field,* Vol. VIII (Accessed 20 March 2012). URL : http://champpenal.revues.org/7971.

Eterno, J., and Silverman, E. 2012. *The Crime Numbers Game: Management by Manipulation.* NYC, CRC Press.

Foucault, M. 2009. *Security, Territory, Population: Lectures at the Collège de France 1977—1978,* New York City Palgrave McMillan.

Foucault, M. 2008. *The Birth of Biopolitics: Lectures at the College De France, 1978-1979,* New York City, Palgrave McMillan

Hammer, M. and J. Champy. 1993. *Reengineering the Corporation: A Manifesto for Business Revolution.* New York City, Harper and Collins.

Maillard, J. d. and T. L. Goff (2009). *"La tolérance zéro en France. Succès d'un slogan, illusion d'un transfert."* Revue française de science politique 59(4): 655-679.

Salmon, Ch. 2007. *Storytelling.* Paris, La découverte.

Sassen, S. 2007. *A Sociology of Globalization.* New York City, Norton and Co.

Silverman, E. B. 1999. *NYPD Battles Crime. Innovative Strategies in Policing.* Boston MA, Northwestern University Press.

Vitale, A., 2008. *City of Disorder. How the Quality of Life Campaign Transformed NY politics.* New York, NYU Press.

Wilson, J. Q. and G. L. Kelling. 1982. "Broken Windows." *Atlantic Monthly* March: 68-71.

crowds & collectivities in networked electoral politics

What happens when a crowd decides to think for itself? **Daniel Kreiss** explores the answer in the 2008 Obama campaign.

President Barack Obama checks his BlackBerry.
(Photo: The White House.)

THROUGHOUT the long primary season, the 2008 Obama campaign's powerful suite of networked electoral tools fueled the candidate's outsider bid for the presidency. A thirteen million member email list helped raise hundreds of millions of dollars online. The nearly two million supporters with active accounts on the campaign's electoral platform helped power the incredible 'crowdsourcing' of field operations. Millions of supporters working alone, or gathering in geographic and affinity groups, canvassed communities across the country, providing critical organizational infrastructure for the campaign in key states.

Supporter involvement was by design in both a planned and technical sense. Staffers in the campaign's New Media Division carefully set their goals, defined metrics for supporter involvement, and then measured their progress. Staffers honed data and analytic practices to optimize the content and format of emails, increase the likelihood of desired supporter actions, and track response rates. Meanwhile, the My.BarackObama.com (MyBO) electoral platform was the product of years of strategic planning and technical development. Consultants designed the platform to facilitate supporter participation in electoral activities such as fundraising, voter identification, and turnout operations. Through these designed-in "affordances" (Norman 1988), these tools translated the extraordinary mobilization around Obama into the money, message, and votes the campaign needed. The campaign's effort to crowdsource these electoral tasks generally proceeded smoothly throughout the primaries and general election. The goals of supporters and the campaign were closely aligned around defeating Hillary Clinton and John McCain. With shared goals, the crowd was content to embrace the tasks that staffers set for it through its technologies that coordinated work in lieu of formal management structures. Like many other crowdsourced efforts in domains from the design of T-

shirts on sites such as Threadless (Brabham 2010) to reviewing patent applications on Peer-to-Patent (Noveck 2006), the goals were clear and outcomes defined (Fish et al., 2011) for this project.

And yet, technologies are always "interpretively flexible" (Orlikowski, 1992), usable for ends their designers did not intend. At extraordinary moments a sizable portion of the crowd was no longer content to have its participation defined by the campaign. Such was the case in the protests that emerged in opposition to Obama's newfound support of retroactive immunity for the telecommunications firms that assisted the Bush administration in its warrantless wiretapping program once he became the nominee. Through their engagement with the campaign's network tools, members of the crowd drew on their latent agency to become a collective and define their own goals, purpose, and identity. In doing so, they leveraged the affordances of the campaign's very own technologies to take collective action and hold the candidate accountable for his policy shift.

THE ROUTINE WORK OF CROWDS

Mobilization presupposes 'crowds.' There must be some motivation, desire, or interest that mobilizes individuals before an entity such as a political campaign can crowdsource a set of actions it needs accomplished. Michael Slaby, the 2008 campaign's Chief Technology Officer, describes how his colleagues "didn't have to generate desire very often. We had to capture and empower interest and desire... We made intelligent decisions that kept it growing but I don't think anybody can really claim we started something."[1] The "transformational and transcendent desire" of Obama's supporters for a new kind of politics (Knorr Cetina, 2009) and the political opportunity to elect a Democrat and African-American to the presidency provided this mobilization. The challenge for

the campaign lay in coordinating these mobilized supporters to act in concert with one another for the staples of contemporary American electioneering: fundraising, voter identification, and turnout.

Obama's staffers faced what sociologist Katherine Chen (2009) describes as the dilemma of "under- and overorganizing" that is a feature of voluntaristic organizations. As Chen argues, the challenge of voluntarism is in crafting hybrid organizational forms that mix "collectivist and bureaucratic practices, but avoid exercising coercive control" (21). Under- and overorganizing cause what Albert Hirschman (1970) referred to as "exit." In voluntaristic relations, individuals can always drop out or leave the crowd, so they need to feel continually that they have agency and that their contributions are meaningful. Too little in the way of formal management means that contributions are not coordinated or useful to the campaign, leading to a highly dispersed crowd pursuing ill-defined and often demoralizing and ineffective tasks. Too much in the way of formal management, such as explicitly dictating what supporters need to do or using the campaign's email list as if it were an ATM, also causes volunteers to drift away. The challenge for staffers lay in transforming amorphous desire into the distinct and recognizable shape of a crowd pursuing defined tasks; and to do so without engaging in the explicit managerial practices that volunteers, particularly those gathering in a highly distributed fashion online, would not accept. At the same time, staffers continually had to invite participation and make it meaningful, but set clear limits around what would be useful to the campaign. As one staffer described, it was always a "delicate dance" to "make sure people feel like they are involved in the campaign without giving them a sense that they are actually setting strategy."[2]

Staffers crafted a number of rhetorical, organizing, and technical

practices that negotiated this dilemma on a daily basis and proved markedly durable for much of the campaign. The explicit design of the campaign's tools helped solve this dilemma of under- and overorganizing. When programmers designed the affordances of the electoral platform, they created a management structure that directed supporter participation. The general invisibility of these design decisions to the crowd allowed the campaign to direct users subtly, avoiding the perception of overorganizing. In this sense, staffers "delegated" (Latour 1992) much coordination work to their tools, which stood in for a formal managerial relationship between the campaign and its volunteer crowds. The campaign's new media tools had designed-in affordances that made it easy for supporters to raise and donate money for the candidate and call voters to identify their preferences. For example, My.BarackObama.com featured fundraising pages that supporters personalized by setting individual goals and circulating appeals to family and friends to donate, much like the paper-based 'walkathon' pledges common in nonprofit fundraising. The platform featured a calling tool that enabled supporters to access targeted scripts for contacting voters and the campaign's voter file online so they could generate and record data on the electorate.

What these tools did not do, however, was crowdsource the candidate's policy platform or electoral strategy. As one highly-active technical volunteer on the Obama campaign describes: "They [Obama's campaign staff] made it very clear that they wanted you to do what you can do to become your own organizer, to use the tools to organize locally in whatever way you wanted to. They set the parameters for what the community could do as well – more specifically in language and in the actual tools."[3] These technical parameters helped ensure that supporter contributions were gen-

1. Slaby, Michael. Interview by author. New York, NY, August 18, 2010.

2. 2008 Obama campaign New Media Division staffer, personal communication

3. Neil Jensen, personal communication, November 10, 2008

erating the resources the campaign needed.

But these tools were not all powerful. While staffers used tools to capture and guide the energy of the crowd, keeping supporters working on garnering electoral resources, there were limits to their managerial authority. The technologies of the campaign were complex organizational and technical achievements predicated on years of development, bodies of knowledge and skill, and the work of dozens of staffers and supporting systems – and they often broke down (Kreiss, 2012). Even more, as Beniger (1986) reminds us, control is not deterministic, but probabilistic. These tools increased the likelihood of individual supporters pursuing the actions the campaign desired, but they did not determine that this would, in fact, be the case. There were also times when there were ruptures in the alignment of the goals and expectations of supporters and the campaign. When these happened, some members of the crowd fashioned themselves into a collective and determined their own ends for political participation. This collective found that the campaign's electoral platform itself could serve as a powerful tool for dissent.

FROM THE CROWD TO THE COLLECTIVE

Obama's crowd generally accepted the conditions of its participation, so long as the goals of the individuals taking distributed action and those of the campaign were aligned. This was the norm for most of the campaign. Supporters expected the campaign to do everything it could to win, wanting staffers to maximize their fiscal and volunteer contributions and extend the candidate's base of support. Supporters wanted Obama's opponents to be defeated and were generally willing to serve in that effort as best they could.

But what happened when the basic terms of participation in the crowd become contested? Despite staffers' best efforts, there were extraordinary moments when a crisis caused fissures in goal alignment among the campaign and its online crowd of supporters. The Get FISA (Foreign Intelligence Surveillance Act) Right protests around the campaign represent precisely such a moment. Upset with the campaign for its abrupt policy shift on FISA during the general election and demanding accountability, some members of the larger crowd of Obama supporters developed an independent sense of themselves and defined their own purpose for action. New media tools, especially the campaign's own, provided "network forums" (Turner, 2006) that facilitated a transition from a crowd of individuals into a collective. These tools allowed aggrieved supporters to find and make themselves visible to one another, hone a sense of collective identity, plan and coordinate challenges to the campaign, and take the fundraising and publicity actions that drew attention to the protest. The collective used My.BarackObama.com, sites such as Facebook and YouTube, and blogs to gather and garner resources and press attention in an attempt to get Obama to change his newfound position. The organizational affordances of the tools the collective appropriated enabled the Get FISA Right protests to scale rapidly. The group on My.BarackObama.com alone swelled to over 15,000 members. These supporters raised enough money to produce a YouTube video and national cable television advertisement.

While staffers never had to respond to the crowd (they simply had to prevent exit), the collective made demands on the candidate that in the end the campaign was forced to address given the scale of the protests. Obama did not change his position on the bill, but the candidate addressed the Get FISA Right supporters directly through a written statement explaining his position. The campaign also made two senior foreign policy advisors available for an online discussion. Dealing with this newly independent collective was frustrating for many staffers, who saw this organizing as distracting supporters from their electoral work. Yet, there was also the acknowledgment that the very tools the campaign deployed and leveraged to get supporters involved had also facilitated this independent supporter action.

In the end, the moment passed and the campaign's response defused the Get FISA Right collective. While there are a number of interpretations of these events, on one level they reveal how the collective had agency vis-a-vis the campaign, much more so than the silent crowd. On another level, however, it reveals the limited degree to which supporters had a voice in the legislative platform of the campaign. And, with little in the way of a viable alternative candidate, over time members of the collective again came to accept the terms of participation and melted back into the crowd.

CONCLUSION

Crowds work towards ends defined by the organizations that convene and gather them as long as the goals for collective action are generally aligned and tasks are rewarding and meaningful. In these routine times, the crowd actively consents to its work being directed towards particular ends; indeed, there is always preexisting mobilization that calls it into existence. The Obama campaign used networked tools to provide its crowd with shape and direction, an arrangement that was remarkably stable for much of the election given the desire to defeat Obama's opponents and the coordinating work of staffers. This alignment and stabilty was a social and technical accomplishment that had to be realized daily.

The crowd always has latent agency, the potential to develop an autonomous sense of itself and act independently. When goal alignment breaks down members of the

crowd can fashion themselves into a collective. This became clear the moment Obama shifted his stance on FISA, as individuals frustrated by the candidate's opportunism sought to act towards non-defined purposes and ends. Networked technologies played a central role in supporting the constitution of this collective and providing it with the capacity to act on a significant scale. In this context, networked tools provided places for activists to gather, create a shared identity, and take action against the campaign based on their grievances. The campaign's own flexible tools helped this collective scale rapidly and take symbolic and social action that attracted press attention, garnered resources, and ultimately forced a response from the campaign. This response, however, reflected the electoral context. The collective knew that exiting the Obama effort would contribute to McCain's candidacy. After eight years of Bush, Obama's supporters were not about to derail the campaign. Staffers ultimately knew that the collective had little bargaining power, and there was little risk in the candidate's minimal response. ☐

DANIEL KREISS *is an Assistant Professor at the School of Journalism and Mass Communication, University of North Carolina at Chapel Hill.*

WORKS CITED

Brabham, Daren. 2010. "Moving the Crowd at Threadless." *Information, Communication, and Society* 13(8): 1122-1145.

Chen, Katherine K. 2009. *Enabling Creative Chaos: The Organization Behind the Burning Man Event.* Chicago: University of Chicago Press.

Fish, Adam, Kelty, Christopher, Murillo, Luis F.R., Nguyen, Lilly, and Panofsky, Aaron. 2010. "Birds of the Internet: Towards a field guide to the organization and governance of participation." *Journal of Cultural Economy* 4(2): 157-187.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States.* Cambridge, MA: Harvard University Press.

Kreiss, Daniel. 2012. *Taking Our Country Back: The Crafting of Networked Politics From Howard Dean to Barack Obama.* New York: Oxford University Press.

Latour, Bruno. 1992. "Where are the Missing Masses? The Sociology of a Few Mundane Artifacts." In *Shaping Technology*, edited by Weibe Bijker and John Law, eds., 225-258. Cambridge, MA, MIT Press.

Norman, Donald. 1998. *The Psychology of Everyday Things.* New York, Basic Books.

Noveck, Beth. 2006. "'Peer to patent': collective intelligence, open review, and patent reform." *Harvard Journal of Law & Technology*, 20(1): 123–262.

Orlikowski, Wanda J. 1992. The Duality of Technology: Rethinking the Concept of Technology in Organizations. *Organization Science* 3(3): 398-427.

Turner, Fred. 2006. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism.* Chicago, IL: University of Chicago Press.

WIAZ
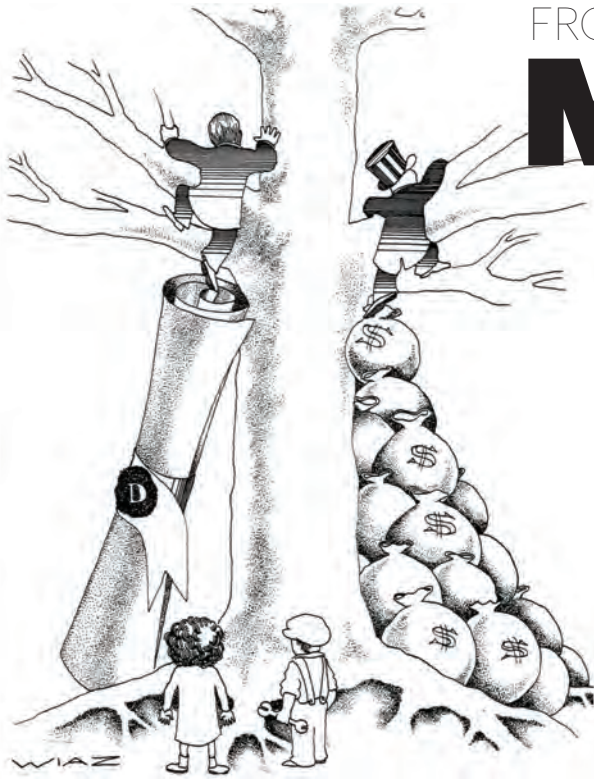
MAPPING THE SOCIAL WORLD

Can data be liberal or conservative? **Alain Desrosières** excavates the curious story of 'correspondence analysis' and its rise to fame.

FOR A LONG TIME, statistics has had a reputation for wiping out individuality, for describing aggregates only through sums and averages. However, in the 1960s, especially in France, techniques of descriptive statistical analysis were developed, (Jean-Paul Benzecri's correspondence analysis) permitting us to focus on individuals within the larger totality. Since then, developments in information technology and the proliferation of quasi-automatic recording (wanted or unwanted) of data from individuals have resulted in techniques known as *data-mining* and *profiling;* shining a light on individuals, for example, to identify future delinquents or simply profile consumers. Statistics has thus constructed an all the more dense network of relations that permits us to make connections between individuals and larger aggregates. Among these mathematical and statistical tools, those most widespread and widely taught are known for their ability to estimate size, draw inferences, and test hypotheses, due notably to the calculation of probabilities. Here we present an altogether different group of tools, more descriptive than inferential, the benefit of which includes the possibility of creating a back and forth between individuals and synthetic representations – obviously something of great interest to political scientists and sociologists.[1]

Belgian astronomer Adolphe Quetelet (1796-1874) introduced to the human sciences the idea of the *average* man, of the regularity and predictability of average behaviors, as opposed to individual behaviors, which are random and especially unpredictable. When human traits, such as size, become "normally" distributed, say according to a bell curve, their average supposedly represents a superior ontological reality, a whole comprised of specific properties, distinct individual cells. This idea would be the basis of future quantitative social sciences, Emile Durkheim's *Le Suicide* being the prototype: sociology is not the uniting of individual psychologies.

Then, at the end of the 19th century, biometricians (and eugenicists) Francis Galton (1822-1911) and Karl Pearson (1857-1936), who touted the idea of inherited biological and intellectual human traits, became interested not only in averages, but also in differences – in the dispersion and distribution of said traits. The individual was indirectly reintroduced, permitting us to locate her along the scale, in space, all while explaining the notion of "correlation" between these traits and the notion of "regression"—formalizing the effects of one

1. The following is part of a more detailed article, dedicated to the history and use of correspondence analysis and published in 2008 in *Journal électronique d'histoire des probabilités et des statistiques (JEHPS)* : "Analyse des données et sciences humaines : comment cartographier le monde social," http://www.jehps.net/Decembre2008/Desrosieres.pdf. An English presentation of Benzecri's methods and application in the social sciences can be found in: Greenacre, Michael and Blasius, Jörg (eds), *Correspondence analysis in the Social Sciences,* Academic Press, San Diego, 1994.

"variable" upon another. But, by focusing on distributions rather than on averages, these new tools introduced the idea of "variation" and eventually "explained variation," and in doing so the individual was temporarily back in a trap. Yet these notions of correlation and regression, the foundations of statistical mathematics, had a very promising future in econometrics, social sciences, and social engineering.

Psychologists Alfred Binet, Charles Spearman, and Louis Léon Thurstone would revive these tools in order to evaluate individuals within larger and more complex spaces through the concept of general intelligence. The factorial analysis of psychologists (principal component analysis), makes visible such multidimensional spaces, but until the 1960s, it was used in psychology much more often than in the other social sciences (political science, sociology, economics).

DATA ANALYSIS À LA FRANÇAISE...
At that time in France a new, multidimensional, analytical tool was put in place by a unique and charismatic statistician, Jean-Paul Benzecri. Called correspondence analysis, it quickly met with success among French sociologists, notably Pierre Bourdieu. It was such a hit because the "fields" of Bourdieu's theory could be represented on cards – graphics with a maximum amount of information contained in a table with a multitude of lines and columns. The cards might note the relative positions of specific individuals or the centers of gravity for clouds of points corresponding to a specific category, for example the employers and bishops in Bourdieu's famous articles. What's more, they're called "dual" analyses: they can show "points/lines" or "points/columns" simultaneously, so that in a single glance, one can see the relative positions of individuals or groups, as well as the variables they represent.

This French-style correspondence analysis was thought by some to be a child of May '68. As it spread through the social sciences around 1970, it was considered "leftist" just as econometric techniques, one the other hand, were thought of as

"rightist." Today this seems strange: statistical tools by themselves are neither "leftist" nor "rightist." How can one explain this phenomenon, typical of the atmosphere following 1968? The arguments (certainly passionately debated) advanced by the tenants of "leftist" data analysis were of two sorts. On the one hand, the analysis was supposedly neutral, with no ideological bias. On the other hand, it was multidimensional.

First and foremost, correspondence analysis was seen as a purely descriptive technique (unlike econometrics), with no underlying, implicit economic/theoretical model: free of the ideologically slanted neo-classical theory. It was supposed to permit one to separate – without any a priori ideology - the fundamental structures buried in an opaque mountain of data. Many said that the idea, fueled by Benzecri, was simplistic because the choice of variables and nomenclatures used in the analytical tables already implied a hypothesis, if not a model. Still, the tool was presented in these terms by its supporters, in explicit opposition to Popperian epistemology, as descriptive exploratory analysis rather than causal analysis bolstered by a predetermined model.

Furthermore, in the wake of 1968, its multidimensionality seemed to be proof of pluralism and democracy, and not simply one-dimensional and reductive (the famous wage scale so dear to economists) – the latter two both synonyms for monotony and hierarchy. Herbert Marcuse's precisely titled *One-Dimensional Man,* one of the epoch's cult books, appeared in 1968 with its vigorous criticism of consumer capitalism. This multidimensionality allowed an understanding of class conflicts more subtle than the proletarian/bourgeoisie split, all while still maintaining the central character of the latter.

The adversaries of this viewpoint traditionally point out that technical tools are without political or ideological tendency, and that mathematical formalism (diagonalization of variance-covariance matrices, eigenvalue and eigenvector search-

es) is the same for correspondence analysis as well as for the solving of econometric models with simultaneous equations. But even if the mathematical *syntaxes* of these two tools are related, their *semantics* are as different as one can imagine: on the one hand, there's the sociological critique of Bourdieu, and on the other, there are the econometric models of government advisors, which focus on action and decision.

...SETS THE STAGE FOR SOCIAL A BI-DIMENSIONAL CARTOGRAPHY...
Bourdieu and his disciples put all this to spectacular use starting in 1975, especially in the book *La distinction: Critique sociale du jugement.* In it, Bourdieu analyzed the tastes and cultural behaviors of the French according to an elegant nomenclature of "socio-professional" groups, which included over thirty positions and which had been used by l'INSEE (French Statistical Institute) since the 1950s for its censuses and research. The interest of this list was that it produced much more complex distinctions than those obtainable with the one-dimensional scale of Anglo-American sociology of yore (upper-class, middle-class, lower-class). Correspondence analysis produced graphic representations structured along "factorial axes" created from data research tables (or matrices).[2]

The first axis, taking into account a maximum of information (or "variance") contained in this table, pitted, as expected, the leisure classes against the lower classes. But the second axis (orthogonal to the first and retaining all the variance not explained by it) showed contrasts that were much less evident a priori (yet with a notably inferior "explained variance," thus maintaining the hegemony of the opposition of the lower-class). Along this second axis, two other categories were contrasted *in probability,* to use Bourdieu's words.

---

2. Technically, these axes correspond to the eigenvectors resulting from the diagonalization of the variance/covariance matrix of the analyzed data table. The explained variances are proportional to the eigenvalues resulting from this diagonalization (Greenacre and Blasius, 1994).

Figure 11 (shown in black)
Variants of the dominant taste. Analysis of correspondences. Plane diagram of 1st and 2nd axes of inertia: the space of properties.

Figure 12 (shown in grey)
Variants of the dominant taste. Analysis of correspondences. Plane diagram of 1st and 2nd axes of inertia: the space of individuals of the different fractions.

The items corresponding to the strongest absolute contributions are indicated in **BOLD CAPITALS** for the 1st factor, CAPITALS for the 2nd.

DISTINCTION
Illustration (with explanations) from Routledge's 2002 edition of *Distinction* by Pierre Bourdieu.

One category was designated "cultural capital" (teachers, artists, researchers, salaried public employees with degrees – mostly urban), and the other was called "economic capital" (employers, merchants, artisans, farmers, salaried workers in private enterprises – more often rural). Thus, at a given level in the Anglo-American scale (for example "the middle-class"), clear distinctions arise in terms of cultural practices, residential neighborhoods, and voting patterns.

The mapping of the social world proposed by Bourdieu in *La distinction* comes in part from correspondence analyses done on data from a number of statistical investigations. Its second axis, contrasting portions of class in terms of cultural capital and economic capital, is relatively stable. It has been upheld by various other studies on consumer practices, on marriage, on the distribution of residential territory in big cities, and on voting patterns. This last example is very telling, for only the bi-dimensional representation allows an accounting of the surprising differences between on the one hand, results of presidential and parliamentary elections, and on the other, votes on the referendum for the European Union in 1992 and 2005.

...THAT MAKES EVIDENT CERTAIN PECULIARITIES IN ELECTORAL BEHAVIOR. In 1970, Benzecri had himself applied correspondence analysis to the voting results of twenty Parisian arrondissements in the 1969 presidential election. Two classic right-wing candidates, Georges Pompidou and Alain Poher, faced off against a communist, Jacques Duclos, two representatives from the intellectual left, Michel Rocard and Alain Krivine, and a candidate who represented small business owners, Louis Ducatel. The first axis classically set Pompidou and Poher against Duclos, the bourgeois neighborhoods against the working class boroughs.

But the statistician Benzecri, in delicious fashion, anticipating what would later constitute the second axis in Bourdieu's celebrated schema in *La Distinction,* commented thusly on the results of the second axis of his correspondence analysis:

On the second axis, we believe it is possible to recognize some common distinctions. Politically, it's Rocard, supported by middle-class intellectuals living in the 6th arrondissement, against Ducatel, whose fiefdom is comprised of another middle-class of artisans and small merchants active between the former Les Halles and the Bastille. On one side are neighborhoods that, while not strictly residential, still possess few workshops and businesses; on the other side, a picturesque maze straight out of Hausmann leading perhaps all the way to Rungis... On the political map, arrondissements 5, 6, 13, 14 and 15 are alone above the first axis, with

Rocard and Krivine... (Benzecri, 1970)

The same bi-dimensional mapping of social categories shows itself to be relevant to the interpretation of differences between traditional political elections and the two referendums concerning the European Union in 1992 and 2005. A map of the results of the 1973 legislative elections, done with surveys indicating the social category of the voters, was published in 1975. Following the first axis from top to bottom (presented vertically), there are five parties. Left-wing voters are cleverly represented on the left of the schema, and right-wing voters to the right. Independent Republicans (the bourgeois right of Valéry Giscard d'Estaing) are on top, in the zone for liberal professionals and high-ranking executives. Below that, on the same vertical and near the center of the image, is the Center (Christian Democrats). The UDR (the Gaullist party) is at the same level as the Center, but more to the right, near the "non-salaried" pole of the second axis. The Socialist Party (that of François Mitterrand) is lower down and clearly to the left, on the side of the mid-ranking executives, while the Communist Party is even lower down, in the worker's zone.

This configuration is a model of classic electoral sociology. However, the bi-dimensional representation allows for a more subtle analysis. The two big political groups, the right and the left (whose vote was split almost equally between Giscard d'Estaing and Mitterrand in 1974), are split on the schema not with a horizontal line, but by a "second bisector" (NW – SE). The Giscard d'Estaing voters are (in probability) more or less upperclass and non-salaried (merchants, artisans, and employers – the categories designated as economic capital). Mitterrand voters are laborers, employees, and the salaried middleclass, notably teachers (categories designated as cultural capital). The National Front, Jean-Marie Le Pen's populist party, did not yet exist. After becoming a significant political force in 1985, it complicated the schema, finding itself socially more

or less in the southeast quadrant of the graph, in the non-salaried, lower and middle-class zone, with voters situated rurally or in small towns.

Thirty years later, this means of representing social space would allow for illuminating clarifications about the voting process in both European referendums: the 1992 Maastricht Treaty and the 2005 European Constitution Project. These elections left the adherents of old school electoral sociology quite disconcerted, because the social groups were not distributed in the same manner as in traditional elections that pit right against left. For these referendums, it's the "first bisector" (NE-SW), *perpendicular to the previous one,* which distinguishes (in probability) the *yes* and *no* voters. In both referendums, the upper classes, urban university graduates, and salaried public employees voted *yes* more often than the lower classes and the small business owners. This was clear from the surveys done in 1992 and 2005.

Over the last two decades, methods called *data mining* have been developed. They are used primarily in marketing, to differentiate categories of clientele, or to "profile" future delinquents. French data analysis like Benzecri's correspondence analysis is the ancestor of more recent tools, regardless of the fact that Benzecri and Bourdieu surely had no inkling of its future applications in business and policing. These descriptive and classifying methods are different tools in inferential statistics, used especially in econometrics and more generally in hypothetical-deductive scientific procedures. Their flexibility is what makes them unique, permitting a back and forth between individuals and their respective regroupings. This flexibility is clearly of paramount importance in both profiling and in the recent explosion of database use, themselves both byproducts of the new information and communication technology developed since 1990.

The history of statistical methods has always been plagued by a tension between the aims of pure knowledge and social criticism on the one hand, and practical application in

the fields of social governance or commerce on the other. This being said, Benzecri's data analysis and more recent methods of data mining cover the entire spectrum, from the most radical criticism up to and including political and commercial endeavors. It is also another and more serious way to pose that naïve question of the 1970s: is correspondence analysis leftist or rightist? ☐

ALAIN DESROSIÈRES *is a statistician and historian at the Centre Alexandre Koyré, EHESS (Paris). He is the author of* The Politics of Large Numbers: A History of Statistical Reasoning, *Harvard University Press, 2002.*

Translated by Paul Knobloch

How do you turn millions of people into a CPU? **Lilly Irani** unravels the mysteries of "human-as-computation" in Amazon Mechanical Turk.

FROM CROWD LABOR TO CLOUD LABOR
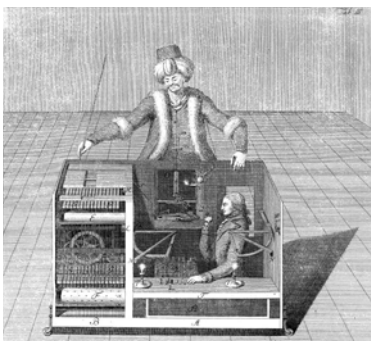
WHEN WE THINK OF computing infrastructure, we think of server farms, personal computers, tangled cables, and operating systems. These are the machines that collect photos, videos, songs, and stories through ubiquitous technologies like Gmail and Facebook. As such data amasses in the wake of web 2.0, technologists have found limits to how far computers using artificial intelligence (AI) can organize and discriminate among such culturally significant data. It is trivial for a person to locate a puppy among a horde of cats and slightly more difficult to guess at family resemblances in a reunion photo, but training computers to perform such feats of cultural discrimination remains an open research problem. Twenty-five years ago Winograd and Flores (1986) declared such problems philosophically insurmountable for AI.

In recent years, however, technologists have found a new workaround to the limits of AI. The "human computation" movement in computer science has advocated for "leveraging the abilities of an unprecedented number of people via the web to perform complex computation" (Law & von Ahn, 2011: viii). The fruits of this research are familiar to anyone who has tried to log into a website only to be challenged with a distorted image of text. Website developers use those images, called CAPTCHAs, to discriminate real people trying to log into a site from password-guessing algorithms trying to break in; CAPTCHA stands for Completely Automated Public Turing test to Tell Computers and Humans Apart. CAPTCHAs succeed at blocking automated break-in attempts building on the observation that recognizing warped text is very hard for a computer but very easy for a literate human being. The more general desire to leverage these computers' and humans' differential capabilities are the foundation of the micro-task marketplace called Amazon Mechanical Turk (AMT).

The design of AMT emerged in the crucible of Amazon's own computational demands for its business. Facing a website riddled with duplicate pages for single products, Amazon engineers had declared using artificial intelligence approaches "insurmountable" (Harinarayan, 2007). Instead, Amazon turned to human computers. To save the time and expense of hiring and managing large numbers of temporary work-

ers, Amazon engineers instead developed a website through which people could work simultaneously from their own computers to check each product for duplicates. Like home-based pieceworkers, these checkers were paid per product evaluated (Pontin, 2007). Workers completed tasks simple and unambiguous enough (ideally) to be completed without coworkers or direct managers; employers paid per unit produced and treated workers interchangeably. In all, the system amounted to a market for largely invisible cognitive pieceworkers. Keeping workers at a distance – here, by mediating them through anonymizing spreadsheets and Application Program Interfaces (API) – allowed Amazon to retain its existing divisions of labor and organizational practices. Where Amazon might have called AI through computer code, now they could call the labor pool similarly. Amazon's CEO publicly announced AMT at MIT saying. "You've heard of software-as-a-service; well, this is human-as-a-service." AMT preserved the social order of technologist-controlled computing, but enhanced this computing with human cognitive capabilities. Six years after Amazon made AMT available for public use, thousands of people do tasks on the service. Computer scientists at MIT and Berkeley have active projects developing databases, word processing tools, task design systems, and other complementary technologies building out the ecosystem of human computing.

So how do the thousands of people unacquainted with one another become a computing crowd? How does AMT extract "ground truth" data from situated cultural cognitions? How does AMT integrate potentially unruly masses into existing large-scale computing infrastructures? I will describe the various lightweight forms of probabilistic control that make AMT work and distinguish them from other highly-controlled computerized workplaces. This analysis builds on three years of experience with AMT through a combination of my role as a builder and maintainer of the AMT worker tool Turkopticon, informal interviews

with technologist employers, attendance at crowdsourcing conferences, and participation in worker web forums.

CROWD CONTROL: ACCOMPLISHING ACCURACY, SPEED, AND SCALABILITY

Accuracy: Technology builders privilege accuracy in the world of AMT. There are two varieties by which requesters try make workers accurate. The first is a sort of statistical objectivity; given the same question, accuracy means exhibiting "the most plural judgement," in the words of then Director of Amazon Web Services Peter Cohen (Sadun, 2006). This can mean simply assigning several workers the same task and using majority vote to decide on the "true" answer, called "the gold standard," or "ground truth" in Computer Science research. More complex mechanisms might try to take into account biasing parameters of the workers such as experience or location. In the end, however, requesters count the most plural as the most accurate and reward workers accordingly. AMT's version of statistical objectivity is a shift in artificial intelligence and natural language processing research, which has traditionally used experts to authoritatively establish "gold standard" data sets (Snow et al., 2008).

The second form of accuracy inheres in tasks that involve subjective or personal data, such as surveys or aesthetic judgments. Requesters need to figure out which workers are making good faith judgments and which ones are "malicious," clicking randomly for money, or trying to corrupt the dataset. AMT maintains an "acceptance rate" for each worker to help requesters recruit workers with high rates of task acceptance. However, large scale requesters use a number of other methods to discriminate "good faith" workers from the "malicious." Most methods boil down to asking obvious questions or providing tasks for which a gold standard is already known. Large-scale requesters maintain databases, organized by alphanumeric ID, recording workers' past performance, geolocation, and other parameters

to create blacklists or whitelists of workers.

One requester I interviewed, for example, put up a digital version of the game Mastermind as a task and found that it was a slightly better predictor of his workers' accuracy than Amazon's reported accuracy rate. But logical acuity is not the only relevant performance. Requesters often restrict workers country location as a proxy for filtering workers without the presumed-to-be-stable cultural literacies their subjective tasks require.

Accuracy is achieved not by training, disciplining, or surveilling workers, but instead by what the founder of one crowdsourcing firm, Rick (a pseudonym), calls "pure approaches" to crowdsourcing. Pure approaches open the system to all workers and use filtering and redundancy to sort the good workers from the unintelligible or malicious, the ground truth from the inaccurate, and the usable data from the spam.

Speed: A hallmark of how we think about computers today is speed – speed in processing, speed in communication. The speed at which AMT accomplishes low-skill or unspecialized tasks is key to its appeal to requesters who might otherwise bring on temporary workers or interns to do the job. The speed at which someone's tasks can be completed on AMT hinges on the size of the crowd present to take on the tasks. A thousand AMT workers working for a single day can process data far faster than hiring ten temps for a hundred days. While a day is still far slower than the near-instantaneous response times we've come to expect from silicon computers, AMT speed will do when usable computer algorithms don't exist and are difficult to create.

Throwing a large number of brains at a problem means having large numbers of people instantly on call. The reach of the Internet into each of the world's time zones means the sun never sets on Amazon's technology platform. Amazon also possesses a uniquely liquid global currency – Amazon Gift Certificates. Though US and Indian workers can get paid in dollars and

rupees respectively, workers from another hundred countries can redeem their earnings in Amazon credits. The global reach of Amazon's currency and website means that whenever someone places a task or makes a call to the marketplace, workers are there to process the tasks.

Scalability: AMT offers developers scalability – developers can command as much or as little human computation as they want, incurring little to no maintenance costs. This scalability stems in part from Amazon's pricing and payment model. AMT charges users per task, giving employers complete discretion to reject work as unsatisfactory and deny the worker payment. Scalability functions in several ways to consumers of AMT labor. For a large corporation doing machine learning categorization, like eBay or Amazon, they need a workforce that is large enough to quickly categorize major inflows of images in parallel during phases when engineers are focusing on improving system performance. For a small startup, scalability offers the promise of low operating costs when small, without sacrificing the promise to investors that the fledgling company can handle the success of rapid growth. The computational quality of scalability, then, is not only technical but also rhetorical.

Scalability also derives from the legal architecture within which the AMT technology is embedded. Amazon's terms of service are designed to allow requesters to pay for data and nothing more. Workers need to understand American English and have access to a computer and an internet connection, but requesters do not pay to train and maintain employees and infrastructure. They pay only for the data workers produce to their liking, and they can refuse on a whim. Amazon's Terms of Participation define workers as contractors providing "work for hire" at prices independent of minimum wage laws. In intent, "work for hire" laws exempt professional contractors from US labor protections with the assumption that those contractors operate and invest in independent businesses

that provide them with opportunities for profit, judgment, foresight, and risk taking. Employers in the US, however, have long attempted with varying degrees of success to specify home workers, pieceworkers, and other low paid workers as contractors as tactic for reducing labor costs (Felstiner, 2011).

On AMT, workers hand over completed work to employers, along with attendant intellectual property rights, regardless of whether the employer approves the work or chooses to pay. Employers can reject the work at their discretion; Amazon neither provides nor advocates for dispute resolution short of mandatory arbitration. Even tax reporting is essentially optional as long as requesters hire each individual worker for less than $600 a year. Amazon's legal architecture leaves requesters free to focus on eliciting and extracting data accurately, quickly, and in a scalable manner. However, there are challenges to managing variously sized crowds of workers within a relatively fixed size organization. Large-scale requesters facing the challenges of managing these crowds are developing techniques I'm calling automatic management.

AUTOMATIC MANAGEMENT AND POLITICS WITH LARGE NUMBERS

For a small start up, managing a workforce of 60,000 people may seem an insurmountable challenge. Yet this is the challenge faced by large-scale requesters. Requesters building on AMT have developed and are constantly refining techniques to manage this workforce in a computer-automated fashion. For AMT to be scalable, the effort that goes into using AMT – setting up tasks, choosing workers, communicating with workers, and deciding who gets paid and who doesn't – must be manageable for someone who might commission 10,000 workers in the span of a few hours. Dahn Tamir, a large-scale requester, explains:

You cannot spend time exchanging email. The time you spent looking at the email costs more than what you paid them. This has to function

on autopilot as an algorithmic system...and integrated with your business processes.

One practice of automated management is "setting up incentives" so that workers self-select into tasks they are good at and learn to avoid tasks they are bad at. "You have to set up incentives right so everyone is aligned and they do what we want them to do. You do it like that, not by yelling at them," another crowdsourcer explained to me. In practice, "setting up incentives" means denying or reducing payment to those who provide work outputs that do not meet requesters needs. The choice of whether or not to pay is based on assessments of accuracy determined algorithmically and is registered through system calls or a spreadsheet upload to the AMT system.

Large-scale requesters also rely on automated filtering criteria, whether based on Amazon's limited worker information (e.g. task approval percentage) or detailed data they gather by interacting with workers. Workers who are below the requesters preferred threshold are simply never shown the task. Those requesters who have more intricate means of sorting "good" workers from the bad may blacklist the bad or whitelist the good. In either case, workers are sorted solely through their performance in the system. At the scale of workforce and the speed of micro-tasks that characterize AMT, there is little time for discipline and little opportunity to mold workers. Sharon Chiarella, VP of AMT, explained that minimal interaction and monitoring allows for efficient human resources management by reducing the decisions employers have to make while simultaneously ensuring that workers are not discriminated against on the basis of race or gender (Chiarella, 2009). This minimalism differs sharply from the surveillance and control of the panoptic, "informated" workplace more typically described (Head, 2003; Zuboff, 1988). Instead, requesters sort desirable workers through faint signals of mouse clicks, text typed, and other digital traces read closely as potential indicators.

Within this large scale, fast moving, and highly mediated workforce, dispute resolution between workers and employers becomes intractable. Workers can contact the requester through a web form on Mechanical Turk if they are dissatisfied with a rejection; but requesters most commonly do not respond personally and Amazon requires no dispute resolution. Requesters have full discretion in choosing to pay workers or even blocking the worker permanently; by AMT's design logic, dispute resolution does not scale. In AMT's transactional logic of data elicitation and automated management, even dispute messages become informational rather than agonistic. Rick admitted that while dispute resolution on AMT scale is impossible, keeping dispute messages in the system gives his company a valuable signal about their algorithm's performance in managing workers and tasks. Disputes, then, become a signal to optimize automated management systems; in AMT as designed, worker struggle consists simply of exit.

Automatic management techniques in AMT are, in a sense, an automation of human resources departments. Recruitment, interviewing, and selection are replaced by an infrastructure that defines the terms of worker entry, entry, exit, and the production of cognitive commodities. These techniques build on much older forms of Taylorism and scientific management; they circumscribe the scope of workers contributions to the overall product as a way of centralizing process planning and consolidating authority in a managerial class (Head, 2003; Noble, 1977). AMT liberates technologists from disciplining workers face-to-face or negotiating over the best way to pursue a goal; it is continuous with the fragmentation of worker collectivities and the centralization of power. In AMT, the Taylorist manager becomes the computer systems programmer. But instead of the total systems view of the informated workplace (Zuboff 1988) advocated by Scientific Management, AMT is made efficient and pleasurable precisely by what employers do not have to know or think about.

AMT and the workforce that powers it become a pleasurable platform for computational innovation. Wendy Chun calls this "causal pleasure"—the sense of power and control a skilled user feels working on and through an operating system on a computer – a "microworld" (Chun, 2005; Edwards, 1996). "Human-as-a-service" places an assemblage of humans and computers under a technologist's interactive control to inspire the technologist's sense of creativity and exploration. As Tamir puts it: "You can try things…When I was wrong, it really didn't matter. I spent a few bucks. The loss was minimal." Accessing workers through APIs, according to founders of one AMT competitor, is key to enabling software engineers to innovate. AMT mediates access to crowds of workers, global competition keeps price-per-task low, and technologists manage those workers lightly, statistically, and expediently. The result is a stable, reliable and enabling infrastructure. In this system, specific workers' agencies – their wrong answers, their complaints, their unwillingness to take a low price, or their choice to leave the labor pool – are largely irrelevant in the operation of a system that structures work to treat people as fungible cognition. Without holding any particular person in "standing reserve" (Heidegger, 1977), AMT's standing reserve of human cognition is achieved. ☐

LILLY IRANI *is a PhD Candidate in the Department of Informatics at UC Irvine.*

REFERENCES

Chiarella, S. 2009. Personal communication at Mechanical Turk Meetup, June 10. San Francisco, CA.

Chun, W. 2005. "On Software, or the Persistence of Visual Knowledge." *Grey Room*, 18, 26-51.

Edwards, P. 1996. *The Closed World: Computers and the Politics of Discourse in Cold War America.* Cambridge: MIT Press.

Felstiner, A. 2011. "Working the Crowd: Employment and Labor Law in the Crowdsourcing Industry." *Berkeley Journal of Employment & Labor Law,* 32(1), 143-204.

Harinarayan et al. 2007. U.S. Patent No. 7,197,459. Washington, D.C.: U.S.

Head, S. 2003. *The New Ruthless Economy: Work and Power in the Digital Age.* Cambridge: Cambridge University Press.

Heidegger, M. 1977. *The Question Concerning Technology and Other Essays.* New York: Harper & Row, 3-53.

Howe, J. 2009. Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business. New York: Random House.

Gantz , J. 2008. "An Updated Forecast of Worldwide Information Growth Through 2011". IDC. Retrieved 2009-04-20.

Law, E. & von Ahn, L. 2011. *Human Computation.* Morgan & Claypool Publishers.

Noble, D. 1977. *America by Design: Science, Technology, and the Rise of Corporate Capitalism.* Oxford: Oxford University Press.

Pontin, J. 2007. "Artificial Intelligence with Help from Humans." *New York Times.* Mar. 25, 2007. Accessed at http://www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html

Sadun, Erica. 2006. "Earning the Big Bucks (Not!) with Amazon's Mechanical Turk - O'Reilly Mac DevCenter Blog." *MacDevCenter.com* Apr 17. Web. 11 Dec 2009.

Snow, R. et al. 2008. "Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks." *Proceedings of the Conference on Empirical Methods in Natural Language Processing.* 254–263.

Winner, L. 1986. *The Whale and the Reactor: A Search for Limits in the Age of High Technology.* Chicago: University of Chicago Press.

Winograd, T. & Flores, F. 1986. *Understanding Computers & Cognition: A New Foundation for Design.* Norwood, NJ: Addison Wesley.

Zuboff, S. 1988. *In the Age of the Smart Machine: The Future of Work and Power.* Basic Books.