Science, Technology & Human Values

REVIEW ESSAY

Corporate capitalism and the growing power of big data

By Martha Poon

Department of Anthropology, The New School For Social Research /

Data & Society Research Institute

**In review. Please do not circulate.**

**6213 words**

### Corresponding Author:

Martha Poon, Data & Society Research Institute, 36 w 20th Street, 11th floor, New York, NY, 10003, USA. Email: martha@datasociety.net

### Reviewed books

Mayer-Schönberger, Victor and Kenneth Cukier. 2013. *Big Data, A Revolution That Will Transform How We Live, Work, and Think*. p 242. Boston: Houghton Mifflin Harcourt.

Bouk, Dan. 2015. *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*. p 304. Chicago: University of Chicago Press.

McFall, Liz. 2014. *Devising Consumption: Cultural Economies of Insurance, Credit and Spending*. p 212. Milton Park, Oxon: Routledge.

### Essay

The baby looks tender and raw, her skin is paper-thin. Eyelids fused together she appears to be sleeping. Chords attached to her belly monitor her heart rate, blood pressure, temperature and oxygenation. The tube inserted in her mouth gives her nourishment, while the one in her nose keeps her breathing. Cradled in the hands of a man we cannot see lays this tiny creature connected to a battery of machines**[[1]](#endnote-1)**.

The person showing us this photograph is internet scholar Viktor Mayer-Schönberger who repeats a version of the message he and journalist Kenneth Cukier insist upon in their best-selling book, *Big Data* (2013): “Lest there be any doubt: Big data saves lives” (p 61).

I’m not convinced. To get some perspective I call my father’s older sister, a retired neonatologist. Auntie has never heard of big data, but over the course of her career, which began in the 1960s, she cared for thousands of newborns and witnessed the development of continuous medical monitoring for premature infants. “It was a helpful tool,” she tells me. “Before that you had you, yourself and your stethoscope. But you couldn’t be there with every baby all the time” [[2]](#endnote-2).

 I tell my aunt that at Toronto’s Hospital for Sick Children, a team of computer scientists from an institute of technology is trying to capture, store and utilize data from medical monitoring – the data exhaust – to providing physicians with analytic information in real time. In an initial study, publicized by IBM, the Artemis system “captured the data stream from bedside monitors and processed it using algorithms designed to spot the telltale signs of nosocomial infection,” that is, an infection the baby contracts in the hospital (IBM, 2012). A stated goal of the research is to drive down the cost of care by reducing the amount of time that babies stay hospitalized (Shein April 4 2012).

 Auntie glazes over. “Proper hand washing is a big thing,” she mutters. Speaking from the perspective of an experienced clinician, she continues, “I don’t know exactly what this thing is that you’re talking about. But even if the machine predicts a possible event, you have to figure out how to intervene. Even if the machine tells you ‘maybe there will be sepsis’ you have to figure out what kind, which bacteria. You have to figure out how to treat it.”

Before hanging up she recommends a novel called *The Cry and the Covenant* (Thompson 1949), a gruesome tale based on historical fact. In the 1840s, in hospitals across the Austro-Hungarian territory, up to one in four women died of childbed fever. The hero of the story, Dr. Semmelweis, orders his students to scrub their hands with chlorine solution to prevent the vicious transfer of decomposing organic matter to laboring mothers from cadavers, other deceased women, whose bodies were being dissected, right there, on the birthing tables. Semmelweis, whose simple technique saved lives, would be dead himself before scientists like Pasteur, Lister, and Koch developed the tools to identify the precise culprit in puerperal sepsis, a bacterium called *Streptococcus pyogenes*.

In *Big Data*, Mayer-Schönberger and Cukier take the position that we do not need to know the causes of things to significantly change outcomes. They are right. But just because data scientists are not in the business of describing causal structures, doesn’t mean that causal mechanisms suddenly melt away in the world at large. Predictive analytics is a tool for fine-tuning the performance of complex physical environments, like the intensive care unit, that sustains the life of premature infants. It is the technical system’s coordinated ability to react, to circumscribe a range of possible responses, to be the cause of its own outcomes that supports life – not the data.

We are still searching for an accurate description of what computer scientists are doing with machine-generated data, especially the data that purport to track and predict human behavior in real time. This is because the word ‘data’ can refer to a much wider set of practices than the ones related to the machine generated traces that circulate in digital infrastructure. As a result, many researchers are opting to think only about data, without attending to the specific qualities of the heavily engineered, networked systems that should be at the center of today’s political debate.

The historians of science at the Max Plank Institute, for example, take big data as an opportunity to reinforce the importance of “data practices stretching back to the 18th century and earlier” (see Hounshell and Midena October 31 – November 2 2013). Similarly, a federation of STS groups in Europe wishes to rehash the politics of quantification (see Grosman February 8-12 2015). Both of these groups set aside networked information infrastructures and personal computers to reinforce the importance of studying well-worn objects of research. Matt Jones is a refreshing exception among the historians of science. His event at Columbia University focused on computational history (see Jones April 17-18 2014).

The conversation has taken a somewhat different turn in the U.S., where social scientists have been assigned to the position of watchdog, a role that was first carved out during the human genome project in the 1990s (Reardon 2013; Zwart and Nelis 2009). The National Science Foundation is supporting a Council for Big Data, Ethics, and Society “to address issues such as security, privacy, equality, and access” (Data & Society Research Institute 2014), while legal scholars are raising concerns about the social ramifications of having algorithms automate decision-making processes (Barocas, Hood and Ziewitz May 16-17 2013; boyd October 12 2015; Pasquale, Petre and Belair-Gagnon April 2 2016; van Hoboken, Nissenbaum and Zeide February 28 2015; ). But again, while these movements look into best practices and possible harms that may result from using predictive analytics they do not initiate basic research on the contemporary phenomenon invoked by the words big data.

Let’s get back to first principles – What’s old and what’s new about data science? What kinds of outcomes can digital data predict? How are these practices related to other forms of business automation? And how did we get here?

Mayer-Schönberger and Cukier are in the thick of it. “Companies that can situate themselves in the middle of information flows and can collect data will thrive,” they announce (p 192). In the guise of public intellectuals, they tout the scientific virtues of big data for an audience of corporate leaders. The book serves up a message that business people are eager to believe: with big data, truth, progress and the pursuit of profit will finally resonate in perfect harmony.

STS scholars won’t buy into this hype. However, we cannot afford to ignore these ideas either. The political economy of big data cannot be harnessed without the intervention of ambitious and deeply capitalized tech companies. Can social scientists challenge the sales pitch that puts the ideal of universal knowledge in the service of the corporate bottom line? Can we explain how and why automation, economic imperatives and personal lifestyle are becoming so tightly intertwined? I think we can. But before we can address these questions we must sift our literature for the elements that are relevant – not all of them are. And we must carefully seek out the angles we’re missing.

Dan Bouk, a modern U.S. historian working within a relatively new field called the history of capitalism, offers a fresh perspective (Beckert 2014; Hyman 2011; Levy 2012; Ott 2011; see Schuessler 2013)[[3]](#endnote-3). In *How Our Days Became Numbered* (2015), Bouk provides insight into the development of data-driven mechanisms in consumer finance.

Researchers in STS have never really considered the developments of commercially motivated systems of data. They have studied the commercialization of science by universities (Mirowski and Sent 2002; Popp Berman 2012). They have also been attentive to technological developments coming out of commercial laboratories (Hughes 1983; Noble 1979; Riordan and Hoddeson 1998). And when the field intersects with management and organizational studies, it does touch upon business issues (Woolgar, Coopmans and Neyland 2009) as well as consumption (Callon, Méadel and Rabeharisoa 2002; Grandclément 2006). Nevertheless, STS has never considered brute financial gain an important driver of data-driven practices.

Donald Mackenzie’s influential article ‘The Credit Crisis as a Problem in the Sociology of Knowledge’, published in the prestigious *American Journal of Sociology* (2011), is a case in point. In his impressive investigation of the back-office events behind the financial collapse of 2008, Mackenzie discovers that a correlation coefficient 0.3 borrowed from a rating agency model for evaluating corporate debt was slipped into a newer model for consumer mortgages. This is a ‘problem’, he vigorously argues, and not, as it might have been described in other contexts, a practical solution, a moment of interpretive flexibility, or a clever piece of bricolage (De Certeau 1988; Lévi-Strauss 1966; Pinch and Bijker 1984).

I’m left wondering. Why are we so willing to deny the first principle of symmetry (Bijker, Hughes and Pinch 1987; Bloor 1999; Latour 1999; Sismondo 2010) to financial outcomes[[4]](#endnote-4)? Mackenzie’s characterization is tantamount to saying that when a handful of people get extraordinarily rich at the expense of others – as some did during the financial events in and around 2008 (see Lewis 2010) – someone, somewhere, must be doing something wrong. This implies that after fighting tooth and nail for radical contingency of the sciences, STS is still willing to concede that financial activity is governed some kind of fundamental principle of justice. But why would financial outcomes be any less politically constructed than scientific ones? How must the analysis be different if they’re not? (For a similar critique of the anthropology of finance see Poon and Wosnitzer 2012; Roitman 2013).

The historians of capitalism are ready to help us out. Bouk documents how profit-hungry corporations have always thrived on churning human lives into data and numbers. The argument follows fast in the footsteps of Jonathan Levy, another historian of capitalism, whose extraordinary book *Freaks of Fortune* (2012) drew a direct connection between the concept of personal risk and the rise of corporate capitalism in the US.

To emphasize the uniqueness of this approach, I’d like to contrast the historians’ position to a couple of classic texts in the social sciences. Consider Ulrich Beck’s hugely popular thesis, *Risk Society* (1992). Beck’s project was to place risk at the center of a theory of western modernity. He argued that our reflexive awareness of the common risks we face results from a general process social transformation, beginning in Europe, in which local social relationships were disembedded, and then rearticulated, through complex technological arrangements, that, in turn, threaten all of human existence (think nuclear meltdown, industrial pollution, climate change etc.).

Bouk and Levy on the other hand, are not doing social theory. They're crafting detailed narratives about controversial characters plucked from the archives of U.S. history. To understand how risk became a property of the person, which is but one type of risk in contemporary social life, the historians remind us that personal risk was invented as part of the violent uncertainties of a boom and bust society in which the losers of the nineteenth century by far outnumbered the winners (Sandage 2005). The historians situate a cultural conception of risk within a set of morally complex relationships to cope with economic instability that were brokered between individuals and early American financial institutions such as “insurance corporations, savings banks, mortgage-backed securities markets, commodities futures markets, and securities markets” (Levy 2012).

 Consider also Ted Porter’s *Trust in Numbers* (1995). These newer history of personal risk, grounded in U.S. corporate capitalism, are very much at odds with Porter’s persuasive argument that quantification appeals to technocrats because it “minimizes the need for intimate knowledge and personal trust” (p xi). If quantification is essentially an impersonal technology that allows experts to act at a distance (p ix; cited in Bouk, p 56) – that is, to act remotely from a center of calculation (Latour 1987) – then we need a very different story to account for how big data intrudes upon the intimate space of the human body (see, for example, Agre 1994; Koponen 2014). Individualized statements of risk can only occur in real time when people are hard-wired, through devices, into specific material infrastructures (see, for example, Burrington 2015a; Burrington 2015b). The proximity of the machines is never fully acknowledged by general concepts like quantification or calculation. Something distinct is happening when a bedside monitor, attached to the body of a premature infant, sits but a heartbeat away.

Personal devices make the attribution of risk exquisitely specific[[5]](#endnote-5). But as Bouk’s research shows, personal risk management pervaded daily life long before the age of digital mediation. To make his demonstration, he takes us back to an unexpected moment: the dawn of life insurance. *How Our Days Became Numbered* charts a course between the practices developed at the great American life insurance companies, starting in the 1830s, and a culture of personal responsiveness to quantified information that Bouk calls ‘statistical individualism’.[[6]](#endnote-6) For him, Americans’ willingness to self-quantify – think bathroom scales, electronic treadmills, calorie counting or Fitbit – is evidence that financial reasoning from insurance, which first conceived of life expectancy as a risk, has seeped into the most intimate recesses of U.S. cultural experience.

Through life insurance, one aspect of human existence that occurs with statistical regularity – the fact that all people die – was transformed into a financial commodity. In theory, an insurance contract is a rational financial product that yields a greater or lesser return from a common pool to its holder, depending on the course of events. In practice, the implementation of insurance markets is never so straightforward. Bouk insists that the financial marketplace is a powerful medium of U.S. cultural production, but he isn’t telling us a story about financial determinism.

*How Our Days Became Numbered* reveals how incredibly messy it has been for insurance companies to move from statistical analysis to a powerful financial industry. The book’s most important contribution is to draws our attention to the fact that questions of commercial justice are often inextricably bound up in debates about technical practices. In Chapter 1, we meet Dr. Scott Lambert of a company called American Popular Life who tried to apply biometry and medical classifications to refine the forecasts in the mortality tables drawn up by the actuaries. Lambert argued that “fitting each life insurance contract as best as possible to each individual’s projected life span” (p 4) would allow for greater equity in the market because short-lives would pay more than long-livers.

Here is one anchor for the history of big data: when investigative journalist today rightfully challenge the justice of using analytics to differentially price consumer goods like staplers or college review courses (Angwin, Mattu and Larson September 1 2015; Valentino-DeVries, Singer-Vine and Soltani December 24 2012), they are protesting a form of individualized pricing that consumer-oriented financial markets have been imagining for a very long time.

In Chapter 2, Bouk lays a second anchor, by raising the issue of discrimination. He warns of the fatalism that can happen when predictive thinking is taken to an extreme. Some early insurance companies refused to offer policies to African Americans having concluded that higher mortality rates meant ‘the race’ was not meant to survive the end of slavery. When policies were finally offered (Chapter 7), the industry’s classing practices “silently ensured that African Americans were held apart and treated differently, as inferiors” (p 203).

All in the name of sound business.

Here is the underexploited thread of Bouk’s narrative: the insurance industry does not rest upon actuarial science or statistical tables alone. The history of statistical practice can never provide a complete explanation of how this industry works because we also need to investigate how companies grappled to make risk profitable. The histories of scientific and business practices overlap, but they are not identical.

Bouk unpacks profit-seeking mechanisms in detail in his fourth chapter, through the case of Mutual Life of New York (MONY), a behemoth of a company that had underwritten $242 million in life insurance even before the onset of the Gilded Age (p 8).

In 1905, MONY was one of the subjects of NY State’s Armstrong investigation into the misdeeds of insurance companies. The heart of the controversy was a practice called smoothing (the opposite of individualization) in which actuaries averaged data in classes, sacrificing individual differences. As Bouk explains, one way to run an insurance pool was to charge policyholders an overly high premium and then pay them back with cash or additional insurance (p 91): “Smoothing helped insurers set their initial rate, but it also did crucial work in the calculation of dividends” (p 93). The example illustrates a fundamental distinction between old Wall Street’s admiration for “the unfettered play of chance” (p 105) and an alternative, heavily administrative form of corporate finance. What Bouk’s reveals is nothing short of stunning. Life insurance was a corruption of the “literal foundations of American capitalism” (p 105).

If social scientists are going to have a critical perspective on how big data analytics and the internet of things are about to reformat economic life, they must retain two points from this story. Firstly, that *technical accuracy is not the foundation of profitability in data-driven industries*[[7]](#endnote-7). Early insurance companies did not make money because data-crunching actuaries churned out better predictions of how death would happen; they made money by managing and manipulating financial elements like the price of the product, the rate at which reserves accumulated, and the calculation of dividends. It stands to reason that big data companies and their advocates are selling horse feathers when they claim that profit is tied to a more perfect science. In today’s funding environment, more likely route to affluence is a financial coup – an infusion of cash from venture capitalist, a big buyout, or a massive IPO.

The second takeaway message is that *financial activity does not follow a singular logic*. Multiple configurations of risk can be implemented in financial markets that can lead to very different patterns of wealth accumulation and can rewrite the boundaries of inequality again and again. Venture capitalist, bankers, hedge fund managers, shareholders, business executives and so on, do not make their money for the same reasons or in the same way. Our financial system today feels fraught, not because the system is inaccurate or in error from some fixed underlying principle, but because the fierce political battle to define ever new and unfamiliar forms of risk never stops (see Poon 2009; Tett 2009).

I wish Bouk had put more emphasis on the insights from the business side of his story. Instead, he pays tribute to STS, highlighting dissonance among the various forms of scientific reasoning that can be used to attribute risk to human beings (i.e. actuarialism, biometry, hygiene, medicine, mathematical biology, statistics and so on.) Digging into these disputes between equally plausible ways of knowing allows Bouk to raise an ethical objection to a culture of risk management that transforms the uncertainty of being alive into an economic cost. In his eyes, no matter what form of scientific practice is utilized, the act of reducing life to unequivocal numbers is a tragic conceit of corporate capitalism.

Despite his personal ethical stance, Bouk makes clear that statistical individualism is a robust form of cultural practice. Through the modern conception of death – the idea that “death could not only be predicted, it could be controlled and resisted” (Chapter 5, p 115) – the cost/risk nexus has become a part, not only of how we think, but how we live. Bouk pointedly observes that “American’s days became things to be forecast and also to be extended, at the same time” (p 218). It was the insurance companies that first encouraged people to modify their behavior and seize their fate by “losing weight, seeing a doctor, or – or, with no apparent irony – cultivating a worry-free lifestyle” (p 218). To reduce payouts and increase revenue the insurance industry became invested in getting all of us excited about prolonging our lives.

The book closes on an intriguing note. Having carefully excavated the business practices of the insurance industry, Bouk turns our attention to the federal government. “Social Security began numbering American’s days with dollar signs” (p 222), he remarks, and “reduced complex men and women to data that could fit on a few punch cards (p 235) provisioned by none other than corporate computing giant IBM (p 230). In a sweeping turn that comes in the final pages, Bouk suggests that big data is a biopolitical form of power (Foucault 2010), only in this case, one that was founded by nineteenth century corporate America, and was later taken up and pushed forward within public health and the services provided by the welfare state.

Which brings us neatly back the data-streaming platform called Artemis, the neo-natal intensive care unit of the publicly-funded Toronto Sick Kids hospital, IBM’s Watson and the data exhaust pouring out of those medical monitors into the hands of computer scientists…

Except for one small fact.

Big data is not an extension of biopower. It can’t be, because the baby is not the one who responds to the information provided by the computer scientists. *Lest there be any doubt*. It is the neonatal intensive care unit – the clinical apparatus – that is supposed to react to the computer scientists’ real-time data, not the helpless neonate dangling on the edge of survival.

Science studies is not as invested as mainstream sociology in identifying types of power (see Callon and Latour 1981; Pickering 1992). If all knowledge can be reduced to a network of local practice, then why bother (Latour 1987b; Latour and Woolgar 1986 (1979))? The topic of big data and networked corporate engineering demands that STS revisit this question. I say this having spent a decade studying how growth and inequality are hardwired into the structure of the financial system. (For an edited collection of case studies see Poon 2013b.)

What kind of power is big data? In other words, what is the mechanism by which these data can be used to exercise an impact in the world? In the tech sector, this is a multi-billion dollar engineering problem. It is also the necessary starting point of a realistic conversation about what equity and social justice can look like in a consumer economy run through personal devices.

In *Devising Consumption* (2014), historical sociologist Liz McFall shows us that data has long been used to extend services to vulnerable groups of people. Where Bouk is concerned with racial inequality in the US, McFall is interested in how financial products were tailored-made to match the UK’s distinctive class structure. She studies a financial product called industrial assurance, which allowed people living in poverty to save for the decent funeral prescribed by Victorian religious customs. Far from obscure, she tells us, “[t]he scale of this industry meant it was the closet thing to a universal system of financial provision for the poor” (p 1).

At base, McFall’s research question is not so different from the one the tech sector is scrambling to solve today: what are the mechanisms and sources of data that can bring about consumption, at scale, even among those of meager means? The Victorian strategy was a fleet of doorstep agents. McFall explains that “[b]y collecting from door to door, agents became the source of discipline necessary to ensure cash-strapped customers paid the regular, usually weekly premiums necessary to service their policies and loans” (p 22). In addition to premiums, agents also collected data. For each home they recorded “payments received and missed, policy types, check amounts advanced, names, addresses and ages”.

The agents were salesmen, not scientists. To make business decisions on the ground, they were trained to consider any impression of the customer they might have. “Signs of opportunity or of risk might be almost anything,” McFall writes. Borrowing a phrase from pragmatist William James, she demonstrates that financial assurance was built up upon ‘the very dirt of private fact’: The state of the garden, the condition of the shoes, a smell in the house, a letter on the mat” (p 5).

The poor are profoundly silent in history, and their choices are rarely documented or understood (Pemble 25 October 2012). What the historical record does reveal is that social reformers of the Victorian era were outraged by “the percentage of income spent on insurance, even in the poorest households” (p 3). Yet McFall reasons that the overwhelming success of the industry is proof that it was deeply in touch with the aspirations of Britain’s lowest classes. The agents were effective at drawing people into the market, she argues, because they were scrupulously attentive, in their own way and for their own collective purpose, to the details of poverty as lived experience. (On this point, McFall is aligning herself with a rich body of work in economic sociology that puts sentiment, in place of rationality, at the heart of economic action (see, for example, Hoang 2015; Illouz 2007; Latour and Antonin Lépinay 2009; Russell Hochschild 1985; Zelizer 1994).)

Writing in defense of sentiment, McFall’s conclusion is as bold as it is unexpected. She indicts impersonally executed welfare programming and critiques the benign detachment built into the early methodologies of the social sciences. “To the extent that market providers devise techniques of sweeping up the mess of elements that make up public and private moods and then incorporating them in products and marketing platforms that seem already to know all about us,” she asserts, “they have the edge over government schemes designed only to appeal to our reason”(p 173). By researching the mechanisms that produce mass financial markets, McFall gives industry its due respect. Her point is that long before social theorists turned their attention towards social history and put the concept of subjectivity at the center of social thought, corporations were actively meddling in consumer experience and personal sentiment as a matter of course.

McFall is a subtle thinker and we need to be mindful of what she isn’t saying. She is not endorsing the field of behavioral economics (Shull and Zaloom 2011); she does not say that markets are the collective result of individual decision-making. Nor is she doing classic science studies; she is not pursuing a history of actuarial or statistical practices (see, for example, Daston 1988; Didier 2009; Gigerenzer et al. 1989; Porter 1988). Most importantly, she makes no spurious claims that assurance markets were fair or just by comparing them to some exogenous standard of epistemological practice. McFall is a philosophical pragmatist. All she says is that assurance boomed because profit-centered enterprise devised a mechanism that allowed it to insinuate financial products into the lives of the poor.

Who is the agent in this story? Who acts? Why the man from the Pru! As McFall points out, in doorstep credit, which developed out of assurance, “It was agents who were in a position to build the relationships that would allow them to assess the creditworthiness of families, as well as motivate them to repay” (p 84). The entity empowered and reinforced by data is the employee of the insurance company, not the pauper pining for a proper funeral, not the consumer. It was through the industry of assurance, that a dedicated fleet of good, average men could become an observable force of social history and an icon of British culture. Dressed in crisp three-piece suits with a bounce in their step, the agents were the organized charisma of corporate capitalism, a model of mass-market subjectivity, a template of middle class virtue whom the struggling classes could emulate.

*Devising Consumption* is a tender tribute to the UK’s iconic doorstep agents who fanned out across the country to remake the population into paying customers. McFall knows the type well. Her smiling father, Bertie McFall, was one of them.

What distinguishes McFall from other economic sociologists is her skillful use of a concept introduced by Michel Callon. To characterize doorstep agents she calls them ‘market devices’ (Callon, Muniesa and Millo 2007). Like the concept of boundary objects (Star and Griesemer 1989), market devices are entities with an informatic content that coordinate action between parties with distinct interests and subjective positions. Imbued with the properties the Silicon Valley transistor (see Lécuyer and Brock 2010; Riordan and Hoddeson 1998), the device, as theory, acknowledges the wide array of communication systems that are engineered to allow mass markets to operate. A credit-scoring algorithm that segments offers in the digital market place can be analyzed as a market device (Poon 2007). So can an automated mortgage underwriting software system (Poon 2009) or a digital slot machine designed to entice play for as long as possible (Dow Schüll 2014; Poon 2013a).

 Market devices are business administrators. They’re not neutral or objective systems, they’re not meant to be (Poon 2013c). In a computer-mediated economy devices are pieces of software programmed to automate the execution of company policies. Are they responsive? Yes, of course! Though feedback and control, consumer-facing devices establish a direct connection between the firm and the marketplace to allow the firm’s to devise, execute and adapt its prerogative. From this perspective, doorstep agents are but one type of device that could reach into households prior to the invention of post-war communication technology.

While effective, the doorstep system was also an extremely costly system of market administration. McFall notes that when the thrifty government officials nationalized the insurance industry in 1911, “the conventions around actuarial valuation, the measurement of mortality risks, the medical examination and the expertise of commercially trained actuaries were all adopted […] while the role of collecting agents was repeatedly refused” (p 59). And yet, from millions of tiny, repetitive transactions with housewives and laborers, assurance providers had accumulated unprecedented reserves of capital. Despite the cost of operations, the industry grew to be such a significant financial power the UK government would draw from these companies to partially fund itself through two world wars (p 174).

The take home message is that companies create data systems that are designed to enable and empower themselves as business operators (see also Yates 1989; Yates 2005). We see this in the history of capitalism, through Bouk, and in the history of consumption, through McFall. The way things look now, in the near future, the digital equivalent of the doorstep agent will be a data-ravenous app, installed on a personal device that automatically adapts a company’s product offering to the user in real time. McFall’s key insight continues to ring true: “To make markets, marketing requires devices, which, by prompting, announcing, listening, collecting, recording, remembering and transmitting information, produce actionable feedback” (p 18).

Big data is valuable because it plays a crucial role in the new digital economy. It is the agent of corporate capitalism, reconstituted as an inexpensive piece of distributed software that transmits data through the infrastructure of the internet.

Networks of machines do not do the same work as human beings, only more efficiently. There is much more to the story of digital automation than cost savings or scale. What is at stake is the quality of life we will have within the political economic structure that creates networked flows of information and that that connects people to corporate control in a continuous and fundamentally new way (Howard 2015; Zittrain 2009).

We can already feel how a computation-based economy is bending the experience of consumption. Thanks to the open movement (see Söderberg 2008) developers have ways of sharing code that cut the cost of making software applications[[8]](#endnote-8). When downloaded onto personal devices, often as freemium, these software-agents can stick much closer to people than a man in a hat ringing the doorbell. The apps generate machine-readable data that can be stored in remote, energy-intensive super-centers as a proprietary asset or commoditized resource (Burrington 2015b; Ensmenger October 11 2015). These data are the raw fuel that grows a powerful new form of operational infrastructure that companies can use to manage markets: algorithms imbued with techniques from artificial intelligence that learn, recursively, on the job.

What remains too little remarked upon is that the tech-sector is rearticulating market society around a set of operational principles that were developed for managing closed systems (Edwards 1997; Mindell 2015). Through ubiquitous computing and the internet of things, how computers work on the inside is being extrapolated to organize a networked socio-economic system. This is why what philosopher of science Stephanie Dick has already observed with regards to the automation of mathematical proofs applies equally to automation through data: “computers can show us that something is the case, but not why or how it is the case” (Dick 2015). Data are part and parcel of how computing systems function, but data analytics do not describe the process by which these systems generate their outcomes. That’s why data science should be categorized as operational control and not as an exercise in the practice of knowledge.

Recent experience shows us that the operational structures that depend upon big data will generate novel mechanisms of value production that were not anticipated by earlier network theorists (see, for example, Benkler 2006). For example, we should not forget that when insurance agents leveraged customer sentiment they extract as much payment as possible from each and every policyholder. In contrast, modern quantitative finance creates products that predict and exploit rates of non-payment (default risk) that neither agents nor old-school loan sharks would ever have tolerated[[9]](#endnote-9). If you’re skeptical of whether these complexly structured products work or not, just remember how strong the banking sector remains. We are at the dawn, not the end of a process of building economic systems that can be administered through data.

I have written this essay to offer a point of caution to my fellow researchers in STS. We should not assume that commercially generated big data is the type of object our field is accustomed to studying. Even if companies can mine their way into record busting profits, consumer tracking is not a project of knowing about people or an offshoot of the enlightenment sciences. We unwittingly promote an oppressive form of high modernist thinking when we assume that economic productivity and justice are bound to some kind of truth inscribed in data. We neutralize all political negotiation with corporate interests when we ask questions about accuracy, meaning or representation instead of unearthing the asymmetrical structures of a digital economy that is increasingly governed by automated systems of feedback and control.

Big data is a robust apparatus wrought by and for the next phase of corporate capitalism. Its purpose is to submit the gestures of digital life to the repetitive decision-making prerogative of profit-seeking enterprises. The major beneficiaries of this infrastructure are not the industrial corporations of old, but a powerful segment of data-masticating, tech-sector unicorns. Backed by venture capital, these companies are rewriting popular culture with the magic of seamless, responsive, real-time convenience. All the while reducing people to a structural version of the neonate, that plugged-in, precarious and machine-subsistent life form that Mayer-Schönberger and Cukier hold up and celebrate.

Close your eyes, take a deep breath, and hang on to your smart phone.

### ACKNOWLEDGEMENTS

I’d like to thank editor David Ribes for the invitation to write this essay. Many interlocutors have helped me articulate the connection between big data and corporate control, including Solon Barocas, Robyn Caplan, Matt Jones, Jannis Kallinikos, Donald Mackenzie, Alison Powell, Johanna Radin, Nico Tempini, Janet Vertesi, Judy Wajcman, and Seth Young. I had conversations by phone with Liz McFall (October 22, 2014) and Dan Bouk (December 15, 2015) to clarify the technical details of their work. My brief exchange with Viktor Mayer-Schönberger and Kenneth Cukier after their public lecture the London School of Economics on March 26, 2013 is available on the record (see footnote one).

### REFERENCES

Agre, Philip E. 1994. "Surveillance and Capture: Two Models of Privacy." *The Information Society* 10:101-27.

Angwin, Julia, Surya Mattu, and Jeff Larson. September 1 2015. "The Tiger Mom Tax: Asians Are Nearly Twice as Likely to Get a Higher Price from Princeton Review." in *ProPublica*. <https://www.propublica.org/article/asians-nearly-twice-as-likely-to-get-higher-price-from-princeton-review>.

Barocas, Solon , Sophie Hood, and Malte Ziewitz. May 16-17 2013. "Governing Algorithms, A Conference on computation, automation and control." NYU Steinhardt.

Beck, Ulrich. 1992. *Risk Society: Towards a New Modernity*. London: Sage Publications.

Beckert, Sven. 2014. *Empire of Cotton, A Global History*: Knopf Doubleday.

Benkler, Yochai. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*.

Bijker, Wiebe, Thomas P. Hughes, and Trevor Pinch (Eds.). 1987. *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge: MIT Press.

Bloor, David. 1999. "Anti-Latour." *Studies in the History and Philosophy of Science* 30(2):81-112.

Bouk, Dan. 2015. *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*: University of Chicago Press.

boyd, danah. October 12 2015. "What World are We Building? 33rd Annual Everett C. Parker Lecture." United Church of Christ.

Burrington, Ingrid. 2015a. "Where the Cloud Rises From the Sea." in *The Atlantic*. <http://www.theatlantic.com/technology/archive/2015/11/where-the-cloud-rises-from-the-sea/415236/>.

—. 2015b. "Why Amazon's Data Centers Are Hidden in Spy Country." in *The Atlantic*. <http://www.theatlantic.com/technology/archive/2016/01/amazon-web-services-data-center/423147/>.

Callon, Michel, and Bruno Latour. 1981. "Unscrewing the big Leviathan: how actors macro-structure reality and how sociologists help them to do so." Pp. 277-303 in *Advances in social theory and methodology*, edited by Karin D Knorr-Cetina and Aaron Cicourel.

Callon, Michel, Cécile Méadel, and Vololona Rabeharisoa. 2002. "The economy of qualities." *Economy and Society* 31(2):194-217.

Callon, Michel, Fabian Muniesa, and Yuval Millo. 2007. "An introduction to market devices." *The Sociological Review* 55(s2):1-12.

Cambrosio, Alberto, and Keating Peter. 1995. *Exquisite Specificity: The Monocolonal Antibody Revolution*. Oxford: Oxford University Press.

Daston, Lorraine. 1988. *Classical Probability in the Enlightenment*. Princeton: Princeton University Press.

Data & Society Research Institute. 2014. "Council for Big Data, Ethics, and Society " in <http://bdes.datasociety.net/>.

De Certeau, Michel. 1988. *The Practice of Everyday Life*. Berkeley: University of California Press.

Dick, Stephanie. 2015. "That is Not Why." *Unpublished manuscript, Harvard University*.

Didier, Emmanuel. 2009. *En quoi consiste l'Amérique? Les statistiques, le New Deal et la démocratie*. Paris: La découverte.

Dow Schüll, Natasha 2014. *Addiction by Design: Machine Gambling in Las Vegas*: Princeton University Press.

Edwards, Paul. 1997. *The Closed World*. Cambridge: MIT Press.

Ensmenger, Nathan. October 11 2015. "The Environmental History of Computing." Pp. 12. Albuquerque, New Mexico: SIGIS Workshop.

Federal Trade Commission. January 2016. "Big Data: A Tool for inclusion or Exclusion?": Federal Trade Commission.

Foucault, Michel. 2010. *The Birth of Biopolitics: Lectures at the Collège de France, 1978--1979*: Picador.

Gigerenzer, Gerd, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty, and Lorenz Krüger (Eds.). 1989. *The Empire of Chance: How probability changed science and everyday life*. New York: Cambridge University Press.

Grandclément, Catherine. 2006. "Le marketing des similarités: les produits à marque de distributeur." *Reseaux* 24(135-136):221-52.

Grosman, Jeremy. February 8-12 2015. "What happens to the data in 'Big Data' ?". FNRS Belgium, Winter School: New Practices of Quantification.

Hoang, Kimberly Kay 2015. *Dealing in Desire: Asian Ascendancy, Western Decline, and the Hidden Currencies of Global Sex Work*: University of California Press.

Hounshell, Eric, and Daniel Midena. October 31 – November 2 2013. "Historicizing Big Data." Max Planck Institute for the History of Science.

Howard, Philip N. 2015. *Pax Technica: How the Internet of Things May Set Us Free or Lock Us Up*: Yale University Press.

Hughes, Thomas Parke. 1983. *Networks of Power: Electrification in Western Society, 1880–1930*. Baltimore: The Johns Hopkins University Press.

Hyman, Louis. 2011. *Debtor Nation: The History of America in Red Ink*. Princeton: Princeton University Press.

Illouz, Eva. 2007. *Cold Intimacies: The Making of Emotional Capitalism*: Polity.

Jones, Matt. April 17-18 2014. "History of Data, Data in History." Columbia University, Hayman Center for the Humanities.

Koponen, Jarno M. 2014. "Your Algorithmic Self Meets Super-Intelligent AI." in *Techcrunch.com*. <http://techcrunch.com/2015/12/14/your-algorithmic-self-meets-super-intelligent-ai/>.

Latour, Bruno. 1987. "Centers of Calculation." Pp. 215-57 in *Science in Action: How to follow scientists and engineers through society*. Cambridge MA: Harvard University Press.

—. 1999. "For David Bloor... and Beyond: A Reply to David Bloor’s ‘Anti-Latour’." *Studies in the History and Philosophy of Science* 30(1):113-29.

Latour, Bruno, and Vincent Antonin Lépinay. 2009. *The Science of Passionate Interests: An Introduction to Gabriel Tarde's Economic Anthropology*: Prickly Paradigm Press.

Lécuyer, Christophe, and David C. Brock. 2010. *Makers of the Microchip: A Documentary History of Fairchild Semiconductor*: The MIT Press.

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

Levy, Jonathan. 2012. *Freaks of Fortune: The Emerging World of Capitalism and Risk in America*: Harvard University Press.

Lewis, Michael. 2010. *The Bit Short: Inside the Doomsday Machine*. New York: W. W. Norton and Company.

MacKenzie, D. 2010. "The Credit Crisis as a Problem in the Sociology of Knowledge." *Forthcoming manuscript*.

Mayer-Schönberger, Victor, and Kenneth Cukier. 2013. *Big Data, A Revolution That Will Transform How We Live, Work, and Think*: Houghton Mifflin Harcourt.

Mindell, David A. 2015. *Our Robots, Ourselves: Robotics and the Myths of Autonomy*: Penguin Random House.

Mirowski, Philip, and Esther-Mirjam Sent. 2002. *Science Bought and Sold, Essays in the Economics of Science*: University of Chicago Press.

Noble, David. 1979. *America by Design: Science, Technology, and the Rise of Corporate Capitalism*: Oxford University Press.

Ott, Julia C. 2011. *When Wall Street Met Main Street, The Quest for an Investors' Democracy*: Harvard University Press.

Pasquale, Frank, Caitlin Petre, and Valerie Belair-Gagnon. April 2 2016. "Unlocking the Black Box." Yale Law School, Information Society Project.

Pemble, John. 25 October 2012. "So Very Silent." Pp. 10-12 in *London Review of Books*.

Pickering, Andrew (Ed.). 1992. *Science as Practice and Culture*: The University of Chicago Press.

Pinch, Trevor J., and Wiebe E. Bijker. 1984. "The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other." *Social Studies of Science* 14(3):399-441.

Poon, Martha. 2007. "Scorecards as devices for consumer credit: The case of Fair, Isaac & Company Incorporated." *The Sociological Review* 55(s2):284-306.

—. 2009. "From New Deal Institutions to Capital Markets: Commercial consumer risk scores and the making of subprime mortgage finance." *Accounting, Organization and Society* 34:654-74.

—. 2013a. "For Financial Certainty, Try Machine Gambling." *Journal of Cultural Economy* 7(4):516-23.

— (Ed.). 2013b. *In case of crisis break glass*. London School of Economics: Center for the Analysis of Risk and Regulation.

—. 2013c. "Response to Tarleton Gillespie’s ‘The Relevance of Algorithms’." in *Governing Algorithms, A conference on computation, automation and control*. <http://governingalgorithms.org/wp-content/uploads/2013/05/1-response-poon.pdf>.

Poon, Martha, and Robert Wosnitzer. 2012. "Liquidating Corporate America: How financial leverage has changed the fundamental nature of what is valuable." *Journal of Cultural Economy* 5(2):247-55.

Popp Berman, Elizabeth 2012. *Creating the Market University: How Academic Science Became an Economic Engine*: Princeton University Press.

Porter, Theodore. 1988. *The Rise of Statistical Thinking: 1820-1900*. Princeton: Princeton University Press.

—. 1995. *Trust in numbers: the pursuit of objectivity in science and public life*. Published Princeton, N.J.: Princeton University Press.

Reardon, Jenny. 2013. "On the Emergence of Science and Justice." *Science, Technology & Human Values* 38(2):176-200.

Riordan, Michael, and Lillian Hoddeson. 1998. *Crystal Fire: The Birth of the Information Age*: W. W. Norton & Company.

Roitman, Janet. 2013. *Anti-Crisis*: Duke University Press.

Russell Hochschild, Arlie. 1985. *The Managed Heart: Commercialization of Human Feeling*.

Sandage, Scott A. 2005. *Born Losers: A History of Failure in America*: Harvard University Press.

Schuessler, Jennifer. 2013. "In History Departments, It’s Up With Capitalism." in *New York Times*.

Shein, Esther. April 4 2012. "Data analytics driving medical breakthroughs, Big data saves lives." in *Computerworld*.

Shull, Natasha Dow, and Caitlin Zaloom. 2011. "The Shortsighted Brain: Neuroeconomics and the Governance of Choice in Time." *Social Studies of Science* 41(4):515-38.

Sismondo, Sergio. 2010. *An Introduction to Science and Technology Studies*: Wiley-Blackwell.

Söderberg, Johan. 2008. *Hacking Capitalism: The Free and Open Software Movement*. New York: Routledge.

Star, Susan Leigh, and James Griesemer. 1989. "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." *Social Studies of Science* 19(3):387-420.

Tett, Gillian. 2009. *Fool's Gold: How the Bold Dream of a Small Tribe at J.P. Morgan Was Corrupted by Wall Street Greed and Unleashed a Catastrophe*. New York: Free Press.

Thompson, Morton. 1949. *The Cry and the Covenant*: Doubleday.

Valentino-DeVries, Jennifer, Jeremy Singer-Vine, and Ashkan Soltani. December 24 2012. "Websites Vary Prices, Deals Based on Users' Information." in *The Wall Street Journal*. <http://www.wsj.com/articles/SB10001424127887323777204578189391813881534>.

van Hoboken, Joris, Helen Nissenbaum, and Elana Zeide. February 28 2015. "Algorithms and Accountability Conference." NYU Information Law Institute.

Woolgar, Steve, Catelijne Coopmans, and Daniel Neyland. 2009. "Does STS Mean Business?" *Organization* 16(1):5-30.

Yates, JoAnne. 1989. *Control through Communication: The Rise of System in American Management*: Johns Hopkins University Press.

—. 2005. *Structuring the Information Age: Life Insurance and Technology in the Twentieth Century*. Baltimore: Johns Hopkins University Press.

Zelizer, Viviana A. 1994. *The Social Meaning of Money: Pin money, paychecks, poor relief & other currencies*. New York: Basic Books.

Zittrain, Jonathan. 2009. *The Future of the Internet--And How to Stop It*: Yale University Press.

Zwart, Hub, and Annemiek Nelis. 2009. "What is ELSA genomics? Science & Society Series on Convergence Research." *EMBO Reports* 10(6):540-44.

###

### ENDNOTES

1. Viktor Mayer-Schönberger and Kenneth Cukier gave a public lecture at the London School of Economics and Political Science (LSE) on March 26, 2013, available online at: <http://www.lse.ac.uk/management/events/public-lectures/public-lectures-archive/130326-BIG-DATA.aspx>. A version of this argument was presented to the authors during question period [49:30-51:22]. The image of a premature infant in their slides [11:00] belongs to professional photographer, J. Hanshaw, in Alabama. <http://www.jhanshawphotography.org/index.php#mi=2&pt=1&pi=10000&s=0&p=3&a=0&at=0>. [↑](#endnote-ref-1)
2. Dr. Kathleen Poon-Vallet is a retired neo-natologist, who trained in Dublin, Ireland and worked with Kaiser Permanente in Honolulu, Hawaii from 1976 until her retirement in 1996. This conversation from April 22, 2013 was recorded as formal fieldnotes. [↑](#endnote-ref-2)
3. Work in history of capitalism seeks to tell stories about broad economic developments and is written in departments of history. Business history, the more traditional field, has tended to focus on the stories of specific firms and is tied to business schools. [↑](#endnote-ref-3)
4. For newcomers to STS: According to the social construction of technology (SCOT), one of intellectual movements absorbed under the umbrella of STS, the same kinds of social and political maneuvers are at work when technologies fail as when they take hold. This implies that the success and failure of technical systems is never predetermined. The outcome depends on the course of events as they actually happen. [↑](#endnote-ref-4)
5. I am borrowing this term from Cambrosio & Keating 1995. [↑](#endnote-ref-5)
6. Bouk refers to ‘statistical individuals’ (noun) but I prefer the term ‘statistical individualism’ (practice) which implies that systemic infrastructure is shaping personal experience. [↑](#endnote-ref-6)
7. There is a longstanding assumption in economic thinking that representational accuracy of data is directly tied to profitmaking. For an example of how the argument plays out in contemporary debate see Federal Trade Commissioner Maureen Ohlhausen’s dissent to the recent FTC report ‘Big Data: A tool for inclusion or exclusion?’ (2016). [Add ref: Donald Mackenzie 2016 on in HFT - ATD becomes profitable by partnering with an investment bank. It is not profitable because algorithmic trading immediately improves the firm’s predictive capacities.] [↑](#endnote-ref-7)
8. Application program interfaces (APIs) are routines, protocols and tools that can be assembled like building blocks. They are used by start-ups to create applications that incorporate and connect to existing platforms like YouTube, Amazon or Google Maps. [↑](#endnote-ref-8)
9. Generating revenue through consistent repayment is a very different beast from a process that uses predictive modeling to manufacture and distributing risk. This is why risk in insurance and risk in consumer credit have distinct genealogies. Life assurance was initially a play on the rate of death over the population, while securitized consumer credit, the preferred model of contemporary global finance, is a play on reducing the administrative cost of controlling repayment while massively increasing credit volume through risk transfer and predictive automation. With securitized credit consumers are routinely offered complex subprime or high-default high-return products without any follow up (Poon 2009). It’s worth nothing that consumers continue express sentiment as they engage with the mortgage industry, but inside financial industry where paper is pooled, traded and valued in complex ways, consumer sentiment is not necessarily the key to extracting value from financial assets. [↑](#endnote-ref-9)