Defining the Sensor Society

Television & New Media
2015, Vol. 16(1) 19–36

DOI: 10.1177/1527476414541552
tvnm.sagepub.com
SAGE

Mark Andrejevic[1] and Mark Burdon[2]

Abstract

The proliferation of embedded and distributed sensors marks the increasing passive-ication of interactivity. Devices such as smart phones, cameras, drones, and a growing array of environmental sensors (both fixed and mobile) and interactive online platforms have come to permeate daily life in technologically equipped societies. Consequently, we are witnessing a shift from targeted, purposeful, and discrete forms of information collection to always-on, ubiquitous, opportunistic ever-expanding forms of data capture. The increased use of sensors marks important changes to our understandings of surveillance, information processing, and privacy. In this article, we explore the transformations associated with the emerging sensing environment. The notion of a sensor society provides a conceptual basis for understanding the characteristics of emerging forms of monitoring and control.

Keywords

data mining, surveillance, privacy, power, sensors, smart phones

From Interactivity to Sensing

A top sales executive at the Ford Motor Company caused a stir at Las Vegas's highly publicized annual Consumer Electronics Show in 2014 when he announced that, thanks to embedded devices in his company's cars, "We know everyone who breaks the law; we know when you're doing it . . . We have GPS in your car, so we know what you're doing" (Edwards 2014, para. 3). Although he later qualified that claim with the assurance that the data are only used with customer "approval or consent" (presumably via a lengthy and obscure "terms of use" agreement), he highlighted an important aspect of a growing array of networked digital devices: they passively collect enormous amounts of data that have wide-ranging potential applications in realms from

[1]Pomona College, Claremont, CA, USA
[2]University of Queensland, Brisbane, Australia

Corresponding Author:
Mark Andrejevic, Department of Media Studies, Pomona College, 333 North College Way, Claremont, CA, 91711, USA.
Email: mark.andrejevic@pomona.edu

marketing to law enforcement and beyond (Sparkes 2014, para. 2). Automobile insurance companies are already using "black boxes" that track driving habits in exchange for discounted rates: "Drive 'well' and you'll keep your discount. Drive poorly and you could see it disappear" (Cooper 2012, para. 4). One marketing company has installed a different type of "black box" in businesses throughout downtown Toronto that tracks mobile phones via the unique identification they send to Wi-Fi networks. The result is that, unbeknownst to the phones' owners, their shopping patterns, dining preferences, and clubbing habits are collected, stored, and shared with participating businesses: "The company's dense network of sensors can track any phone that has Wi-Fi turned on, enabling the company to build profiles of consumers' lifestyles" (Dwoskin 2014, B1).

These are just two examples of the ways in which forms of pervasive, always-on, passive information collection are coming to characterize the use of digital devices and the business models with which they are associated. If, once upon a time, the mobilization of the promise of interactivity was characterized by the enthusiastic portrayal of heightened forms of active participation on the part of users, the automated collection of data "passive-izes" this interactivity. These days, we *generate* more than we participate—and even our participation generates further and increasingly comprehensive "meta"-data about itself. Our cars, phones, laptops, Global Positioning System (GPS) devices, and so on allow for the comprehensive capture of the data trails users leave as they go about the course of their daily lives. In the business world, this device-driven data—combined with new techniques for putting it to use—have been enthusiastically greeted as a valuable economic resource: described as the "new oil," it is treated as a resource to be extracted, refined, and put to use (Deutscher 2013, para. 3). The familiar moniker of "big data" is a direct result of proliferating forms of "interactive" data capture because it refers to the burgeoning reserves of data generated by a growing array of sensors and made available for various forms of sorting, sharing, and data mining.

In this regard, the rise of "big data," the fascination with the figure of the "data scientist," the development of new forms of data analytics, and the "passive-ication" of interactivity are interlinked via increasingly powerful and comprehensive sensing devices and networks. We propose the concept of the "sensor society" as a useful way of approaching these interconnections and exploring their societal significance. The term is meant, in the first instance, to refer to a world in which the interactive devices and applications that populate the digital information environment come to double as sensors. In many instances, the sensing function eclipses the "inter*active*" function in terms of the sheer amount of information generated. For example, the amount of data that a smart phone generates about its user in a given day is likely to far surpass the amount of data actively communicated by its user in the form of text messages, e-mails, and phone calls (not least because each of these activities generates further data about itself: where the text was sent, how long the call lasted, which websites were visited, and on and on).

But the notion of a "sensor society" also refers to emerging practices of data collection and use that complicate and reconfigure received categories of privacy,

surveillance, and sense-making. Finally, the notion of the sensor society is meant to direct attention toward the costly infrastructures that enable data collection, storage, and processing as well as to the advantages that flow to the institutions that own, operate, and access them. There are structural asymmetries built into the very notion of a sensor society insofar as the forms of actionable information it generates are shaped and controlled by those who have access to the sensing and analytical infrastructure. Some of the main attributes of an emerging sensor society include the following: the increasing deployment of interactive, networked devices as sensors; the resulting explosion in the volume of sensor-generated data; the consequent development and application of data mining and machine learning techniques to handle the huge amounts of data; and the ongoing development of collection, storage, and analytical infrastructures devoted to putting to use the sensor-derived data.

Viewed through the lens of the "sensor" society, conceptions of interactivity and notions of privacy and power appear in a somewhat different light than in recent celebrations and critiques of digital media. Database-generated forms of "knowledge" that are "too big to know" (Weinberger 2011, 1) are not accessible in the way that other forms of knowledge are. As we shall argue, data mining privileges those with access to the data and the technology when it comes to generating actionable information that may be neither fully explicable (in the sense of being illuminated by an underlying explanation) nor reverse-engineerable. In the following sections, we consider in greater detail the significance of these characteristics of the emerging sensor society and their implications for new forms of data collection, monitoring, and surveillance.

The Rise of Sensors

Any networked interactive device can double as a sensor insofar as it collects and relays data about how it is used, and these data can be used to infer information about the user and/or the user's environment. For a smart phone, for example, to provide accurate and continuous location awareness, the device has to connect to a variety of local Wi-Fi access points (or cellular network towers) while in transit. The transmission of these data not only enables the device's functionality, but it also means that the device can double as a sensor, and there are a growing range of apps that can be used to collect data about users and their activities (Dwoskin 2014, paras. 1ff). This logic is generalizable across the digital landscape: devices and applications developed for one purpose generate information that can be repurposed indefinitely. For example, the scanners that allow cashiers to enter prices more rapidly can also be used to track the speed at which employees work; digital video recorders capture data about viewing habits (including paused and fast-forwarded moments); e-readers capture data about when and where a book is read, which passages or pages are skipped; Facebook recently announced a mobile app that uses the microphones in smart phones to detect nearby music or TV shows (Makarechi 2014); and so on.

Sensing technologies and apps for the smart phone industry alone have spawned a rapidly expanding market as new sensing frontiers unfold. For example, the U.S. Department of Homeland Security has funded a program to develop smart phone

sensors that can detect toxic chemicals in the air to provide an early warning system for industrial accidents or terrorist attacks. Smart phone users would, in effect, become distributed mobile sensors automatically relaying data back to the Department of Homeland Security (DHS 2013) about air quality. By the same token, employers increasingly rely on a range of sensors to monitor workers: keystroke monitoring software, smart cards that track employee movements, GPS devices that monitor drivers and delivery personnel, and even applications that track employees' facial expressions (Waber 2013).

Researchers at MIT have even developed wearable monitoring devices called "sociometers" that automatically track "the amount of face-to-face interaction, conversational time, physical proximity to other people, and physical activity levels" among workers to "measure individual and collective patterns of behavior, predict human behavior from unconscious social signals, identify social affinity among individuals working in the same team, and enhance social interactions" (MIT Media Laboratory 2011). Even employee recruitment practices are being sensorised. A company called Evolv that mines large sets of recruitment and workplace data reported as one of its key findings that "people who fill out online job applications using Web browsers that did not come with the computer . . . but had to be deliberately installed (like Firefox or Google's Chrome) perform better and change jobs less often" (*The Economist* 2013). The web browser used to fill out a job application becomes an important element of the job application itself. As such examples indicate, the Internet provides a model for the sensor society, insofar as its version of interactivity is one in which, increasingly, movement through cyberspace generates data that can be collected, stored, and sorted. Digital sensors form an interactive overlay on the physical spaces they populate, allowing them to become as trackable as the Internet. Thus, devices like Google Glass, for example, transpose the affordances of cyberspace (back) into the register of physical space: locations can be tagged and book-marked.

As such applications proliferate, our devices and our environments are likely to become increasingly populated by sensors in what would once have seemed surprising ways: car seats with heart-rate monitors, desks with thermal sensors, phones with air quality monitors, tablets that track our moods, and so on. Once information about our mood inferred through our facial expressions, body temperature, pulse, and so on can be collected, a new array of sensors can be developed to respond to these data—and, in turn, to collect, store, and make sense of the data generated by this response. When interactive devices are treated as sensors, creative uses for existing data sets can be developed and new sensing capabilities can be piggy-backed upon existing ones. Consider, for example, the efforts of Microsoft researchers to develop apps that transform smart phones into "mood sensors" (LiKimWa 2012, 1). Rather than developing a specific biometric sensor to detect mood (via, say, electroencephalogram [EEG] readings, skin conductance, voice stress, etc.), the researchers simply tracked the ways in which users' self-reported moods correlated with their usage patterns, and then developed a model that built on these findings to predict mood, reportedly with 94 percent accuracy (LiKimWa 2012, 23ff). As new forms of sensing and data collection are devised, these are leveraged against already existing data troves that have

accumulated over years. The sensor-derived data and its collection can be repurposed indefinitely.

In this regard, sensor-derived data collection is dissimilar to traditional practices of surveillance even though sensor-related collection activities trigger similar surveillance and monitoring concerns. In their report on "The Surveillance Society" for the U.K. Information Commissioner, David Murakami Wood and Kirstie Ball propose a preliminary definition of surveillance as "purposeful, routine, systematic and focused attention paid to personal details, for the sake of control, entitlement, management, influence, or protection" (Wood and Ball 2006, 1). They further emphasize that "surveillance is also systematic; it is planned and carried out according to a schedule that is rational, not merely random" (Wood and Ball 2006, 3). Similarly, in his influential formulation of "dataveillance," Roger Clarke refers to "the systematic monitoring of people or groups, by means of personal data systems, in order to regulate or govern their behaviour" (Clarke 2003, para. 5). Clarke subsequently distinguished between targeted personal dataveillance and "mass dataveillance, which involves monitoring large groups of people" (Clarke 2003, para. 7). Although the forms of sensor-based monitoring associated with interactive media technologies share broadly in these logics of information collection, they also differ in important ways.

If, for example, as Wood and Ball (2006) argue, surveillance is focused and in reference to identifiable persons, this is only partially true of sensor-based forms of monitoring. The goal of sensor-related collection is the capture of a comprehensive portrait of a particular population, environment, or ecosystem (broadly construed). More systematic forms of targeting start to take place against this background, and increasingly come to rely on it. The population-level portrait allows particular targets to emerge—and once they do, their activities can be situated in the context of an ever-expanding network of behaviors and the patterns these generate. Thus, sensor-derived surveillance can be untargeted, non-systematic, and often opportunistic. Consider, for example, the fact that some U.S. military drones are equipped with a device called an "Air Handler" that can capture all available wireless data traffic in the area through which the drone flies. As one of the rare news accounts about this device put it, when a drone goes out on a mission, "the NSA [National Security Agency] has put a device on it that is not actually under the control of the CIA or the military; it is just sucking up data for the NSA" (Goodman 2014). The drone then comes to represent a double-image of surveillance: both the familiar "legacy" version of targeted, purposeful spying and the emerging model of increasingly ubiquitous, opportunistic data capture. As one news account puts it, "the NSA just wants all the data. They want to suck it up on an industrial scale. So they're just piggybacking on these targeted operations in an effort to just suck up data throughout the world" (Goodman 2014, para. 8). For drones, the signal-saturated sky is a sea of electromagnetically stored data that can be scooped up, processed, refined, and perhaps put to use.

Such examples highlight the additive, convergent, and intersectional character of surveillance associated with sensor-based data acquisition. As new sensors come online, the data they capture can be added to existing databases to generate new patterns of correlation. The goal is not necessarily to follow or track an individual, per se,

but to capture a specific dimension of activity or behavior across the interactive, monitored space—to open up new data-collection frontiers (mood, gait, typing patterns, preferred browser, etc.) in an expanding "digital enclosure," wherein a growing range of spaces, places, and the information associated with them enter into the monitored embrace of digital interactivity (Andrejevic 2007). This type of data capture gives new meaning to the notion of focused monitoring: not exercised upon a particular individual per se but upon a specific dimension or register of activity. Even if individuals are not the target of the pattern generation process, it becomes easier than ever before to identify them, sort them, and target them. New sensors open up new dimensions of the population, environment, or ecosystem. Once these dimensions are developed, they can be compared with others to generate potentially useful patterns for purposes ranging across a range of activities from politics and policing to health care, employment, education, marketing, and more. The goal is to broaden the range of monitored dimensions that give shape to the population–environment nexus, allowing it to emerge in new ways as a site of detection, measurement, analysis, and intervention.

Defining the Sensor Society

Concepts such as "the information society" (Beniger 1986; Webster 2007, among others) and "the surveillance society" (Lyon 2001, among others) have relatively broad currency in both the media studies literature and popular media discourses, so what justification might there be for yet another sweeping moniker? The notion of a "sensor society" clearly not only fits within these broader categories, but it also isolates a salient aspect of emerging social logics so as to focus attention upon them and their broader implications for social, cultural, economic, and political life. The notion of a "sensor society" (Schermer 2008), then, is meant to focus attention on developments in the collection and use of information in the digital era that might help re-orient discussions about issues ranging from surveillance and power to privacy and social sorting. The frame of the "sensor-society" addresses the shifts that take place when the once relatively exceptional and discrete character of monitoring becomes the rule, and when the monitoring infrastructure allows for passive, distributed, always-on data collection. Our hope is that directing attention to the logic of sensing-based monitoring will open avenues for further exploration of the dimensions of a sensor society in which the devices we use to work and to play, to access information and to communicate with one another, come to double as probes that capture the rhythms of the daily lives of persons, things, environments, and their interactions.

In general terms, a sensor is a device that measures or detects an event (such as an earth tremor or a status update) or state (such as the temperature) and translates this measurement or detection into a signal: it "responds to stimuli" (the "sensitive element") and "generates processable outputs" (the "transducer") that are translated into "readable signals" by a "data acquisition system" (Kalantar-Zadeh and Wlodarski 2013, 12–22). To view a device as a sensor within the context of the sensor society is to approach it from a particular angle: to determine what type of information the sensor automatically collects (what it measures or detects), how this information is stored and shared, and how it can be put to use.

Sensors can include any device that automatically captures and records data that can then be transmitted, stored, and analyzed. A keystroke monitoring system on a computer that can record the unique speed and pattern of an individual's typing style is a form of sensor, as is a web browser that can capture and record someone's Internet search habits (it detects and transduces). These devices may be much more than sensors, but they partake of the logic of sensing as a form of passive monitoring, and can be treated as, among other things, components of an increasingly comprehensive, albeit distributed and often disarticulated sensing apparatus. Some sensors may be coordinated with others, but others rely on infrastructures that are owned and operated by distinct entities that do not necessarily share information with one another. Sensors do not watch and listen so much as they detect and record. They do not rely on direct and conscious registration on the part of those being monitored. When one sends an e-mail to someone, one is actively communicating to them, but when a device detects the details of one's online activity (including e-mails), sensor-based monitoring is taking place.

Thus, new realms of interactivity open up new dimension of sensing and intervention, as do new technologies and practices. When automated license plate readers and radio frequency identification (RFID) scanners were developed, it became possible to trace mobility in new ways. When phones went mobile, they traced new frontiers in geo-locational monitoring. These monitoring dimensions are further expanded by the addition of Internet access and other interactive applications. A dedicated sensor is not necessary to expand the sensing frontier: thanks to data mining techniques, e-mail, phone activity, or browsing behavior can turn personal devices into mood detectors, illness monitors, and fitness evaluators. We might divide these developments up into new technological frontiers in sensing (the development of new forms of dedicated sensors—location tracking devices, expression detectors, infrared or ultrasound detectors, toxic chemical detectors, etc.) and expanding frontiers in *inferential sensing* (the ability to extrapolate information from the data provided by the existing technology and dedicated sensors—such as inferring mood based on texting and web browsing activity). In this sense, the data mining process helps to expand the available dimensions of sensing.

The Explosion of Sensor-Derived Data

The shift away from targeted to comprehensive forms of data collection may be enabled by new, inexpensive, and distributed forms of networked devices; but it is driven by the logic of emerging forms of data analysis. When the goal is to generate as much data as possible to discern otherwise inaccessible and undiscernible patterns, it is impossible to determine in advance the full range of potentially useful types of information. The goal then becomes to collect as much data as possible in as many dimensions as are available. Unsurprisingly, then, the amount of data collected on a daily basis is historically unprecedented but is, nonetheless, a small foretaste of things to come. IBM (2013, paras. 1ff) claims, for example, that every day, about 2.5 quintillion bytes of data are generated—the data equivalent of a quarter million copies of the

print collection of the Library of Congress—and that 90 percent of the world's stored data have been created in the past two years. That is, if the entirety of recorded human history were shrunk to the length of a day, the vast majority of its accessible stored data would have been created in the equivalent of the last thirty seconds. Much of these data are generated mechanically and automatically by a burgeoning array of digital sensors that capture not just human activity but climate data, traffic flow, machine activity, and so on. However, the upshot is that sensor-derived data accumulate faster than human hands can collect it and faster than human minds can comprehend it.

Capturing, storing, and making sense of huge amounts of data are a resource-intensive endeavor, even despite the falling costs of digital storage and processing power. The costs continue to escalate in part because what counts as "big data" continues to dramatically increase, and in part because the goal of total information capture is built into the data mining model. The Central Intelligence Agency's (CIA) Chief Technology Officer, Gus Hunt, has described this Google-inspired approach as a paradigm shift for intelligence agencies insofar as they are moving "away from search as a paradigm to pre-correlating data in advance to tell us what's going on" (Hunt 2012). *All* data are potentially useful in this framework:

> The value of any piece of information is only known when you can connect it with something else that arrives at a future point in time . . . Since you can't connect dots you don't have, it drives us into a mode of, we fundamentally try to collect everything and hang on to it forever. (Sledge 2013)

The result is that big data mining remains the preserve of large corporations and well-funded agencies. What counts as data about "everything" continues to grow as new forms of sensing and sense-making are developed.

Thus, one of the characteristic challenges for emerging forms of sensor-derived data collection is the sheer amount of information they generate. For example, when the avalanche of images generated by U.S. surveillance drones threatened to outstrip the ability of human observers to make sense of them, out-of-the-box thinkers at the RAND Corporation turned to a seemingly unlikely source for inspiration and assistance: reality TV producers (Menthe et al. 2012). The latter had extensive experience in sorting through hours of uneventful tape to isolate a few decisive moments. The logic uniting drones and reality TV, according to military think tankers, is the need to rapidly process the large amounts of information generated by twenty-four-hour, multi-camera video surveillance. As one news account puts it,

> . . . when you start thinking about some of these reality shows that have dozens of cameras, continuously running, and then these producers trying to compartmentalize all of that and cram it into a 30-minute episode, you start to get an idea of how much they may have in common with the Air Force. (CNN 2012)

The RAND Corporation report is a meditation on the difficulties posed by the human bottleneck in processing the tremendous amounts of data generated by sensors. The

problem is not a new one in the "intelligence" world: signals intelligence in the post–World War II era has long posed the challenge of information glut: how best to make sense of the increasingly large amounts of information that can be captured, stored, and viewed or listened to by intelligence analysts. Recent developments in data mining and analytics indicate that the tendency will be away from human analysis and toward automated forms of data processing.

The CIA's Gus Hunt describes the shift toward data mining as one that replaces the older "search and winnow" model, in which a small portion of useful data is kept and the rest discarded. Thus, the CIA's rationale for sweeping up as much data as possible is representative of the logic permeating predictive analytics: the value of some forms of information is speculative in the sense that it cannot be determined until further "data points" arrive. The very *possibility* of utility warrants collection under conditions in which technological developments make it possible to store more and more data due to the proliferation of sensors and the explosion of sensor-derived data. Given the additive and speculative character of data mining (a data set might yield new and useful patterns when paired with future information), the purpose and justification for monitoring in the sensor society can come after the fact.

Meta-datafication

The automated capture and storage of data give rise to another important aspect of this data explosion—what might be described as the process of *meta-datafication*—the treatment of content as just another form of meta-data, or (by the same token), the understanding that the only real content of interest, from a data analytical perspective, is that which is machine-readable. Consider, for example, Google's oft-repeated rejoinder to those who accuse the search-engine giant of disregard for privacy because of its aggressive information collection and tracking practices: "no humans read your email or Google Account information" (Byers 2013). Machines do not attempt to *understand* content in the way a human reader might. Instead, they scan e-mail and online behavior for potentially useful patterns. The substance of this rejoinder to privacy concerns is that people should not worry because Google's machines have transformed the meaningful content of their communications into meta-data: not actual content but information about the content (what words appear in it, when, where, in response to whom, etc.).

It is precisely the potential of the automated processing of sensor-derived data that underwrites the productive promise of data analytics in the sensor society: that the machines can keep up with the huge volumes of information captured by a distributed array of sensing devices. Treating the content of e-mail as meta-data is one of the consequences of transforming networked communication devices into sensors that capture the behaviors and communications of users. Accordingly, one of the lessons of the sensor society is that content can be treated as meta-data, insofar as emphasis on the ideational content is displaced by the focus on patterns of machine-readable data. Perhaps this shift is what MIT's Big Data guru Sandy Pentland is gesturing toward when he claims that

> . . . the power of Big Data is that it is information about people's behavior instead of information about their beliefs . . . It's not about the things you post on Facebook, and it's not about your searches on Google . . . Big data is increasingly about real behavior, and by analyzing this sort of data, scientists can tell an enormous amount about you. (*Edge* 2013)

Pentland's distinction does not hold up: what one posts on Facebook—along with detailed information about when, where, and how—*is* a form of behavior, as are one's search patterns on Google. What Pentland is really getting at is what might be described as the vantage point of "big data," which privileges a perspective that focuses on information as a pattern-generating form of behavior and not as ideational content. Jeremy Packer (2013, 298) sums up this perspective shift in his description of a model, "pioneered and widely implemented by Google" in which, "the logic of computation is coming to dominate. In this model, the only thing that matters are directly measurable results"—what Pentland describes as "behavior." As Packer (2013, 298) puts it,

> Google's computations are not content-oriented in the manner that advertising agencies or critical scholars are. Rather, the effect is the content. The only thing that matters are effects: did someone initiate financial data flows, spend time, consume, click, or conform? Further, the only measurable quantity is digital data. Google doesn't and couldn't measure ideology.

This shift is what Pentland most likely means when he says that Facebook posts and search requests are not of interest. That is, they are not of interest from an *ideational* perspective. As behavior, of course, they help provide valuable data. The messages themselves, when read by the machine, become, in a sense, contextual information *about themselves* (and users) even when they are isolated from the ideational content of the message to a particular receiver.

The notion that the collection of meta-data is somehow less powerful or intrusive than that of the content with which they are associated has come under considerable scrutiny (Narayanan and Shmatikov 2010; Ohm 2010). Former Sun Microsystems engineer Susan Landau, for example, confided to the *New Yorker* magazine that the public "doesn't understand" that meta-data is "much more intrusive than content" (Mayer 2013, para. 22). It is possible to unearth intimate details about individuals without having a human actually read their communications. Knowing where people go at what times of day, whom they communicate with, and so on, can reveal a lot about them, including sensitive information about their health, their political inclinations, and their private lives.

It should come as no surprise that, from a privacy perspective, the process of meta-datafication erodes the concept of information privacy and the laws that flow from that concept. Different definitions exist as to what constitutes personal information, but typically information privacy law deals with information that can be used to identify an individual. Personal information can therefore be specific data or combinations of data that can identify individuals directly, such as full name, driver's licenses, or social

security numbers, but it can also include data that indirectly identifies individuals. For example, a residential address can be used to aggregate different sets of data that facilitate identification. The legal definitions of personal information recognize that the nature of personal information generation is inherently contextual. Information can become personal information in different contexts, at different times, and in different social relationships.

The logic of the sensor society envisions the prospect that individuals will be uniquely identifiable from the meta-data created by sensor devices and sensor networks. That is, seemingly anonymous information such as patterns of movement or online search behavior, or even unique typing patterns, can give rise to the identification of unique individuals, especially in an environment where more and more sensors collect a growing range of data. As data from different sensors is combined and mined, it is possible to infer further information about such individuals—including details that would, in other contexts, fall into protected categories—without needing to know their names or their addresses. However, given the ease with which these data can eventually be traced back to named individuals by drawing upon combinations of databases, *all* data about persons harvested from increasingly comprehensive sensor networks are likely to become, for all practical purposes, personally identifiable.

The Search for Un-anticipatable Patterns

Because the proliferation of sensors underwrites the recent explosion of digitally stored data and pushes necessarily in the direction of automated data processing, the forms of knowledge generated by automated forms of data mining become characteristic of a sensor society. These forms of knowledge rely upon emergent processes in the sense that their goal is to generate un-anticipatable and un-intuitable correlations: that is, patterns that cannot be predicted in advance. Thus, the imperative for more data is not simply a result of the desire to gain as complete a record as possible of populations and environments but also of the data mining process itself: un-anticipated or un-intuitive results can be generated by adding in new categories of data, even seemingly irrelevant data. For example, the fact that browser selection correlates with job performance is not something that employers would be likely to anticipate—it is an artifact of the data mining process, which included a consideration of variables not traditionally associated with job interviews but made available through the mechanics of the online job recruitment process. The data miners used the information because it was available to them—part of the trove of information collected during the application process but not intentionally incorporated into that process. There is a rationale to this kind of monitoring, but it is neither systematic nor targeted. Analysts do not start out with a model of the world that they are setting out to prove or disprove, like a detective trailing suspects, but with a trove of information. This trove is shaped by the available sensing technology, much of which is, in turn, the result of affordances built into devices, networks, and applications for a range of reasons that might initially have little to do with the goals of those who seek to put the data to use.

The Ongoing Quest for Diachronic Omniscience

Predictive analytics and related forms of data mining extend into the future the quest for what Lisa Parks (2005, 91) presciently calls "diachronic omniscience." Parks uses the term to describe the forms of comprehensive information capture associated with satellite-based forms of monitoring: "the capacity of media to comprehensively record global space through time" (as paraphrased by Russill 2013, 102). The hope on the part of data miners is that comprehensive data about what happened can be used to project into the future. However, as Parks has demonstrated in her discussion of satellite imaging, when the data from sensors accumulate, they can be used not only to model the future but also to mine the past. Consider, for example, the use of digital records to link suspects to crime scenes. Police have already used mobile phone data to catch thieves by placing them at the scene of the crime and reconstructing their movements in a subsequent car chase (Perez and Gorman 2013). The goal of "diachronic omniscience" invokes the possibility of a complete archive that could supplement the vagaries of reported actions and memories by externalizing them in the form of machine-readable databases. The related claim to the repeated (but highly contestable) refrain that we need not worry about new forms of data collection as long as we are not doing anything wrong is that the database can provide us with alibis. Alternatively, for those who are guilty, the archive can be used to link them to the scene of a crime, to reconstruct their movements, to identify, and to eventually capture them.

Any attempt to approach so-called "diachronic omniscience" necessarily entails the formation of databases large enough to re-duplicate the world in informational form and the development of analytic tools to make sense of these data. The issue of infrastructure is accordingly central to these examples and thus to any critical consideration of the sensor society.

Jeremy Packer (2013, 297) captures something of this logic in his echo of the Kittlerian call to attend to infrastructure:

> Understanding media not merely as transmitters—the old "mass media" function—but rather as data collectors, storage houses, and processing centers, reorients critical attention toward the epistemological power of media . . . Media forge real power/knowledge relationships that reassemble the world.

By contrast, the airy rhetoric of "cloud computing" and various notions of "immateriality" that have been associated with digital, post-industrial forms of production and consumption represent what might be described as a turn away from infrastructure in both popular and academic discussions of digital, networked media. Not that long ago, brand-name futurists including Esther Dyson and Alvin Toffler proclaimed the "central event of the 20th Century" to be the "overthrow of matter"—and along with it allegedly anachronistic preoccupations with property, hardware, and infrastructure (Dyson et al. 1996, para. 1). Even Hardt and Negri's (2009, 294) conception of "immaterial labor" pushes in the direction of imagining a "self-valorizing" productivity

unfettered from the constraints of fixed capital: "Today, productivity, wealth, and the creation of social surpluses take the form of cooperative interactivity through linguistic, communicational, and affective networks." The tendency of such formulations is to direct attention toward particular types of expressive and communicative activity and away from the often privately owned and opaque infrastructures upon which they rely.

The notion of the sensor society, by contrast, redirects attention toward the infrastructures that make data collection capture, storage, and processing possible and consequently to the relations of ownership and control that shape who has access to data and who sets the parameters and priorities for using that data. Consider, for example, an account of the frustration evinced by one of the generals who helped oversee the development of the Predator drone (one of the more highly publicized technological icons of the sensor society): "he has grown so weary of fascination with the vehicle itself that he's adopted the slogan 'It's about the datalink, stupid'" (Bowden 2013, para. 12). The drone, like the sensors distributed across the networked digital landscape, is "a conduit": "Cut off from its back end, from its satellite links and its data processors, its intelligence analysts and its controller, the drone is as useless as an eyeball disconnected from the brain" (Bowden 2013, para. 12). In other words, the sensor is inextricably related to the communication and analytical infrastructure upon which it relies. Sensors can, of course, operate at close range, such as the devices that detect whether a smart phone is in bright light or close to someone's head. However, it is when these data can be captured, stored, and shared—that is, when the sensors are articulated to the infrastructures for data collection and analysis (and eventual response)—that the salient characteristics of the sensor society emerge.

Making Sense of the Sensor Society

The proliferation of sensors pushes in the direction of automation: not simply in the data collection process but in data analytics and response. Because the sensing process is not discrete, but continuous, and because the target is not a particular individual or moment but what might be described as a defined dimension (and any event that takes place in that dimension), the data accumulates indefinitely. In broader terms, the additive goal behind the proliferation of sensors can be understood to be the digital replication of entire populations and environments enabled by a variety of distinct but interconnected infrastructures. Individual targets are treated as pieces of a puzzle. All of them must be included for the puzzle to be complete, but the picture is not of them or about them, per se, but about the patterns their data form in conjunction with that of others. In the sensor society, the target is the pattern and the pattern is an emergent one (insofar as it cannot be detected until the analysis process is undertaken).

Conventional understandings of privacy as control over one's self-disclosure and self-presentation are complicated by this reconfiguration of targeting toward patterns rather than people and especially by the emergent character of pattern generation. The turn toward automated forms of predictive analytics means that it is, by definition,

impossible to reasonably anticipate the potential uses of the information one discloses. The goal of data mining large quantities of information is, *by definition*, to generate un-anticipatable and un-intuitable predictive patterns (see, for example, Chakrabarti 2009). That is, the data analytic process is systemically and structurally opaque. It follows that data collection and analytical infrastructures are equally opaque. The legal theorist Tal Zarsky (2013, 1519) describes the decisions based on such data mining processes as "non-interpretable" (and thus non-transparent) because of their inherent complexity:

> A non-interpretable process might follow from a data-mining analysis which is not explainable in human language. Here, the software makes its selection decisions based upon multiple variables (even thousands).

As such, processes of opacity that yield un-anticipated uses for data that result in un-interpretable decisions undermine some of the key foundations of information privacy law, namely, informed consent and even ideas such as contextual integrity (Nissenbaum 2010). To the extent that the ongoing generation of un-anticipated uses becomes the norm, the norms lose regulatory purchase: they do not rule out any particular use in advance. The search for unpredictable and otherwise indiscernible correlations means that so-called "function creep" is not ancillary to the data collection process but is built into it: the function *is* the creep. Increasingly, all data need to be treated as personal data in the sensor society because any given piece of data, aggregated with other available databases for the purpose of predictive pattern generation, could have the capacity to identify an individual but more importantly could be used in a way that impacts on their life chances.

Neither the concept of information privacy law nor anti-discrimination law is designed to cope with the vastness of data collection and analysis envisioned by the sensor society. All data simply cannot be personal information under the rubric of information privacy law. All decisions of exclusion cannot be discriminatory under anti-discriminatory law. Quite simply, the legal systems created around these concepts would fail to operate if that was the case. Regulation of the sensor society thus poses a new set of legal challenges (Cohen 2012).

Underwriting these observations is the recognition that the sense-making processes and the sensor technology must be considered in conjunction with one another. The sensor society we are describing is inseparable from both its back-end infrastructure and from the logics of sensor-driven data analysis and response. The ability to collect large amounts of data becomes associated with new forms of sense-making (that rely on capturing as much information as possible and on predicting and correlating rather than explaining or understanding). Big data mining approaches push in the direction of more comprehensive data collection and thus embrace the imperative of developing more comprehensive sensing networks. Thus, the invocation of the notion of a sensor society looks beyond the ephemeral construct of "big data" to invite a critical interrogation of the power structures that shape the development and use of sensing and sense-making infrastructures.

To forward the notion of a "sensor society" is not to posit the wholesale transformation of all forms of information capture, processing, and use. We do not seek to contest critical claims about surveillance in the digital era, so much as to add a further dimension—albeit one that we argue is unique and significant. Nor do we claim to have exhaustively described the sensor society—which is an emerging phenomenon—but we do hope that by defining a particular perspective, we have opened up avenues for further exploration, both conceptual and empirical. Not all the attributes we describe as characteristic of a sensor society are unique to it, and yet, we argue that their combination is unique and significant and that current popular, academic, and regulatory discourses have not yet caught up with them or taken them fully into account. Our hope is that in outlining the notion of a sensor society, we have highlighted some key issues related to surveillance, monitoring, privacy, and control for the foreseeable future. We anticipate that the study of what might be described as the cultural, social, political, economic, and technological logics of the sensor society will become an increasingly pressing concern as interactive devices proliferate and become equipped with a growing array of increasingly powerful sensors. It is the task of those who seek to understand these developments to ensure that their theoretical, conceptual, and critical formulations keep pace with the technology and its deployment.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: One of the authors was supported by a research grant: Australian Research Council Discovery Project Grant (DP1092606).

References

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

Beniger, James R. 1986. *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, MA: Harvard University Press.

Bowden, Mark. 2013. "The Killing Machines: How to Think about Drones." *The Atlantic*, August 14. http://www.theatlantic.com/magazine/archive/2013/09/the-killing-machines-how-to-think-about-drones/309434/ (accessed May 15, 2014).

Byers, Alex. 2013. "Microsoft Hits Google Email Privacy." *Politico.com*, February 7. http://www.politico.com/story/2013/02/microsoft-renews-google-attack-on-email-privacy-87302.html (accessed May 15, 2014).

Chakrabarti, Soumen. 2009. *Data Mining: Know it All*. New York: Morgan Kaufmann.

Clarke, Roger. 2003. "Dataveillance—15 Years On." *Personal Website*. http://www.rogerclarke.com/DV/DVNZ03.html (accessed May 15, 2014).

CNN. 2012. "The Situation Room." September 10. Transcript retrieved online at: http://edition.cnn.com/TRANSCRIPTS/1209/10/sitroom.02.html.

Cohen, Julie. 2012. *Configuring the Networked Self: Law, Code and the Play of Everyday Practice*. New Haven: Yale University Press.

Cooper, Charlie. 2012. "Backseat Big Brother." *The Independent*, August 9. http://www.independent.co.uk/life-style/motoring/features/backseat-big-brother-is-the-insurance-companies-black-box-worth-it-8022694.html (accessed May 15, 2014).

Department of Homeland Security. 2013. "Cell-All: Super Smartphones Sniff Out Suspicious Substances," Official Website. http://www.dhs.gov/cell-all-super-smartphones-sniff-out-suspicious-substances (accessed May 15, 2014).

Deutscher, Maria. 2013. "IBM's CEO Says Big Data Is Like Oil, Enterprises Need Help Extracting the Value." *Silicon Angle*, March 11. http://siliconangle.com/blog/2013/03/11/ibms-ceo-says-big-data-is-like-oil-enterprises-need-help-extracting-the-value/ (accessed May 15, 2014).

Dwoskin, Elizabeth. 2014. "What Secrets Your Phone Is Sharing about You—Businesses Use Sensors to Track Customers, Build Shopper Profiles." *Wall Street Journal*, January 14, B1.

Dyson, Esther, George Gilder, George Keyworth, and Alvin Toffler. 1996. "Cyberspace and the American Dream: A Magna Carta for the Knowledge Age." *The Information Society* 12 (3): 295–308.

Edge. 2013. "Reinventing Society in the Wake of Big Data: A Conversation with Sandy Pentland." August 30, 2012. http://www.edge.org/conversation/reinventing-society-in-the-wake-of-big-data (accessed May 15, 2014).

Edwards, Jim. 2014. "'We Know Everyone Who Breaks the Law' Thanks to Our GPS in Your Car." *Business Insider* (Australia), January 9. http://www.businessinsider.com.au/ford-exec-gps-2014-1 (accessed May 15, 2014).

Goodman, Amy. 2014. "Death by Metadata: Jeremy Scahill and Glenn Greenwald Reveal NSA Role in Assassinations Overseas." *Democracy Now!* (radio program), February 10. Transcript retrieved online at: http://www.democracynow.org/2014/2/10/death_by_metadata_jeremy_scahill_glenn (accessed May 15, 2014).

Hardt, Michael, and Antonio Negri. 2009. *Empire*. Cambridge, MA: Harvard University Press.

Hunt, Gus. 2012. "Big Data: Operational Excellence Ahead in the Cloud," Presentation to the Amazon Web Services Government Summit 2011, Washington, DC, October 26. http://www.youtube.com/watch?v=SkIhHnoPpjA (accessed May 15, 2014).

IBM. 2013. "The IBM Big Data Platform." *IBM Software Group (Web Page)*. http://public.dhe.ibm.com/common/ssi/ecm/en/imb14135usen/IMB14135USEN.PDF (accessed May 15, 2014).

Kalantar-Zadeh, Kourosh, and Wojciech Wlodarski. 2013. *Sensors: An Introductory Course*. New York: Springer.

LiKimWa, Robert. 2012. "MoodScope: Building a Mood Sensor from Smartphone Usage Patterns" (Doctoral dissertation, Rice University, Houston, TX).

Lyon, David. 2001. *Surveillance Society*. Buckingham: Open University Press.

Makarechi, Kia. 2014. "Facebook Knows What Music You're Listening To." *Vanity Fair*, May 22. http://www.vanityfair.com/online/daily/2014/05/facebook-listens-music-tv-shows-share (accessed May 28, 2014).

Mayer, Jane. 2013. "What's the Matter with Metadata." *The New Yorker*, June 6. http://www.newyorker.com/online/blogs/newsdesk/2013/06/verizon-nsa-metadata-surveillance-problem.html (accessed September 2, 2013).

Menthe, Lance, Amado Cordova, Carl Rhodes, Rachel Costello, and Jeffrey Sullivan. 2012. "The Future of Air Force Motion Imagery Exploitation: Lessons from the Commercial

World." *The RAND Corporation: Project Air Force*. http://www.rand.org/content/dam/rand/pubs/technical_reports/2012/RAND_TR1133.pdf (accessed March 15, 2014).
MIT Media Laboratory. 2011. "Sociometric Badges." http://hd.media.mit.edu/badges/ (accessed May 15, 2014).
Narayanan, Arvind, and Vitaly Shmatikov. 2010. "Myths and Fallacies of Personally Identifiable Information." *Communications of the ACM* 53 (6): 24-26.
Nissenbaum, Helen. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford: Stanford Law Books.
Ohm, Paul. 2010. "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization." *UCLA Law Review* 57 (6): 1701–77.
Packer, Jeremy. 2013. "Epistemology Not Ideology OR Why We Need New Germans." *Communication and Critical/Cultural Studies* 10 (2–3): 295–300.
Parks, Lisa. 2005. *Cultures in Orbit: Satellites and the Televisual*. Durham: Duke University Press.
Perez, Evan, and Siobhan Gorman. 2013. "Phones Leave a Telltale Trail." *The Wall Street Journal*, June 15. http://online.wsj.com/article/SB10001424127887324049504578545352803220058.html (accessed September 2, 2013).
Russill, Chris. 2013. "Earth-Observing Media." *Canadian Journal of Communication* 38 (3): 95–116.
Schermer, Bart. 2008. "Privacy and Visibility in the Sensor Society." *SlideShare*. http://www.slideshare.net/Considerati/privacy-and-visibility-in-the-sensor-society (accessed 15 May, 2014).
Sledge, Matt. 2013. "CIA's Gus Hunt on Big Data." *Huffington Post*, March 21. http://www.huffingtonpost.com/2013/03/20/cia-gus-hunt-big-data_n_2917842.html (accessed May 15, 2014).
Sparkes, Matthew. 2014. "Ford Boss Retracts Claim that 'We Know Everyone Who Breaks the Law.'" *The Telegraph*, January 10. http://www.telegraph.co.uk/technology/news/10563828/Ford-boss-retracts-claim-that-we-know-everyone-who-breaks-the-law.html (accessed June 17, 2014).
The Economist. 2013. "Robot Recruiters: How Software Helps Firms Hire Workers More Efficiently." April 6. http://www.economist.com/news/business/21575820-how-software-helps-firms-hire-workers-more-efficiently-robot-recruiters (accessed May 15, 2014).
Waber, Ben. 2013. *People Analytics*. London: FT Press.
Webster, Frank. 2007. *Theories of the Information Society*. London: Routledge.
Weinberger, David. 2011. *Too Big to Know: Rethinking Knowledge Now that the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*. New York: Basic Books.
Wood, David. M., and Kirstie Ball. 2006. "A Report on the Surveillance Society." *Surveillance Studies Network, UK*. http://ico.org.uk/about_us/research/~/media/documents/library/Data_Protection/Practical_application/SURVEILLANCE_SOCIETY_SUMMARY_06.ashx (accessed May 15, 2014).
Zarsky, Tal. 2013. "Transparent Predictions." *University of Illinois Law Review* 2013 (4): 1503–70.

Author Biographies

Mark Andrejevic is an associate professor in the Department of Media Studies, Pomona College. He is the author of *Reality TV: The Work of Being Watched, iSpy: Surveillance and*

Power in the Interactive Era, and *Infoglut: How Too Much Information Is Changing the Way We Think and Know,* as well as articles and book chapters on surveillance, digital media, and popular culture.

Mark Burdon is a lecturer in the TC Beirne, School of Law, the University of Queensland. His primary research interests are privacy law and the regulation of information sharing technologies. He has been a researcher on a diverse range of multi-disciplinary projects involving the reporting of data breaches, e-government information frameworks, consumer protection in e-commerce, and information protection standards for e-courts. His research is published in leading law/technology journals in the United States, the EU, and Australia.