Defining the Sensor Society

Abstract

Sensor technologies are proliferating in our networked environments. Devices such as smart phones, cameras, drones, and a growing array of fixed environmental sensors and interactive online platforms now permeate all aspects of our lives. This developing environment is causing radical changes to traditional forms of information collection, storage and analysis processes. We are witnessing a shift from targeted, purposeful and discrete forms of information collection to always-on, ubiquitous, ever-expanding and non-specific forms of data generation and acquisition. The increased use of sensors therefore marks important changes to our understandings of surveillance, information processing, and privacy. In this paper, we create a new lens that examines the radical information changes unfolding. We label this new lens the sensor society and we provide a conceptual basis to understand four currently distinct attributes, which when put together, provide a different viewpoint of how our technologically driven society is evolving and the potential consequences that are emerging.

1. Introduction

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). Although he later qualified that claim with the assurance that the data is 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 marketing to law enforcement and beyond (Sparkes, 2014). 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). One marketing company has installed a different type of "black box" in businesses throughout downtown Toronto that track mobile phones via the unique identification they send to Wi-Fi networks. The result is that, unbeknownst to the phone's 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).

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-fies" 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, 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 -- has been enthusiastically greeted as a valuable resource: the 'new oil' (Deutscher, 2013). The familiar moniker of "big data" is a direct result of proliferating forms of "interactive" data capture, since it refers to the burgeoning reserves of data available for various forms of sorting, sharing, and "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. We propose the notion of the "sensor society" as a useful way of approaching these inter-connections 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, emails, 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 Web sites 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 even sense-making. The defining attributes of this 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 predictive analytics to handle the huge amounts of data; and the ongoing development of collection, storage and analytical infrastructures devoted to making sense of 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. 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.

2. The Rise of Sensors

Our networked digital devices no longer simply fulfil a specified purpose or range of purposes. They can now act as sensors that generate, detect and collect information about our activities, about our environments and about entire societies. In order 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 this data enables the device's functionality but it also means that the device can now act 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). 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; the phones that people carry with them can collect and rely information about their location, their communication practices, and their contacts; 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, 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 US 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 either industrial accidents or terrorist attacks. Smart phone users would, in effect, become distributed mobile sensors automatically relaying data back to the DHS about air quality (Department of Homeland Security, 2013).

By the same token, employers increasingly rely on a range of sensors to monitor workers: key-stroke monitoring software, smart cards that track employee movements, GPS devices that track drivers and delivery personnel, and so on (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 in order 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 upload a job application becomes an important element of the job application itself. Indeed, 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 space they populate, allowing it 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, etc. Once information about our mood through our facial expressions, body temperature, pulse, and so on can be collected, a new array of sensors can be developed to respond to this data – and, in turn, to collect, store and make sense of the data generated by this response.

When all interactive devices can be 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). Rather than developing a specific biometric sensor to detect mood (via, say, 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, allegedly with 94% accuracy (LiKimWa, 2012). 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 driven data and its collection can be endlessly repurposed.

In this regard, sensor driven data collection is dissimilar to traditional notions of surveillance even though sensor related collection activities bear the same hallmarks of surveillance and monitoring concerns. In their report on "The Surveillance Society" for the UK Information Commissioner, David Murakami Wood et alia 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 et alia, 2006). 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 et alia, 2006). Lyon outlines a comparable concept of the 'surveillance society' that encapsulates the same features (Lyon, 2008). 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). Clarke subsequently distinguished between targeted personal dataveillance and "mass dataveillance, which involves monitoring large groups of people" (Clarke, 2003). While 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.

Sensor-based data generation and collection can be opposed to purposeful, routine, systematic and focused attention. Individuals are not targeted in the conventional sense. Instead, sensors monitor populations and environments. Data analysis of sensor-based data consequently relies on blanket coverage to discern larger, actionable patterns and insights that can then be re-deployed at the individual level.

If, as Wood et alia (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 is amorphous, non-discrete and unspecific.

This highlights the additive, convergent, and intersectional character of surveillance inherent in 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 to follow or track a specific 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 what Andrejevic (2007) has called "the digital enclosure." This type of monitoring 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. New sensors open up new dimensions of the population, environment, or ecosystem. Once these dimensions are developed, they can be compared with other dimensions to generate potentially useful patterns for purposes ranging across the spectrum 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 measurement, analysis, and intervention. Unlike the surveillance society, the purpose and justification for monitoring in the sensor society can come after the fact. With that in mind, we outline the concept of the sensor society and its implications in the following sections.

3. Defining the Sensor Society

Concepts such as "the information society" (Webster, 2007 and Beniger, 1986 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 justification is a familiar but nonetheless compelling one: that isolating a salient aspect of emerging social logics might help focus attention upon them and their broader implications for social, cultural, economic, and political life. The notion of a "sensor society" (Schermer, 2008) is not an exhaustive one, in the sense of attempting to explain all aspects of information use in contemporary forms of social practice, nor is it an exclusionary one in the sense of superseding or replacing other approaches. However, it is meant to focus attention on particular changes 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 control.

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. In the following sections we explore some significant aspects of these shifts in order to trace the outlines of emerging forms of sensing. The discussion is meant to be suggestive and productive rather than exhaustive: further aspects of the sensing paradigm will inevitably emerge as data collection practices develop alongside techniques for information processing. 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 communicate with one another come to double as probes that document the rhythms of the daily lives of persons and things. The logics we explore will, we anticipate, continue to develop as new sensors and new forms of automated sensing and sense-making multiply, capturing data of all kinds that converge in the omnivorous maw of the database.

We start by defining the four attributes of the sensor society and then consider the logics that link them in Section 4.

a) Sensor Characteristics

Our focus so far has been on devices that have the characteristics of a sensor. This raises a key question – when does a device partake in sensing activity? In this sub-section, we outline the characteristics of sensors to define the first attribute of the sensor society, how devices operate as sensors. We do not see this as delineating which devices are sensors and which are not, but rather of defining when and in what capacity a device that may have many functions is acting as a sensor.

In general terms, a sensor is a device that measures or detects something and translates this measurement or detection into a signal: it "responds to stimuli" (the "sensitive element"), "generates processable outputs" (the "transducer") that are translated into "readable signals" by a "data acquisition system" (Wlodarski and Kalantar-Zadehz). 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 analysed. 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 on a smart phone, for example, can detect changes in movement and translate these into signals that can be stored and transmitted (and then sorted and analysed) – allowing, for example, information about one's daily movements or even about the idiosyncrasies of someone's gait to be detected and recorded (Biosensics, 2014). The action of the sensor is automatic and receptive. Sensors don't 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 email to someone one is actively communicating to them, but when a device detects the details of one's online activity (including emails), sensor-style monitoring is taking place.

We are arguing for a particular perspective on interactive devices as sensors that highlights the characteristic forms of monitoring with which they are associated. We might think of the various ways in which sensors and devices that function as sensors are coming to permeate life in technologically developed societies. Once one starts thinking about interactive devices as sensors – a range of possibilities opens up: cars that can detect stress levels (via heart rate monitors, for example) could help insurance companies more accurately assess risk; online movie viewing can provide data about preferences to guide future recommendations as can biometric measurements of affective response, the Google map application on a smart phone can be used to generate real time traffic maps, and gaming platforms can even be used for national security purposes (Kolakowski, 2014).

New realms of interactivity open up new dimension of sensing and intervention, as do new technologies and practices. When automated license-plate and rfid scanners were developed it became possible to trace mobility in new ways alongside the exploitation of new affordances. When phones went mobile, they traced new frontiers in geo-locational monitoring. Ditto when they added internet access and other applications. As we have seen, a dedicated sensor is not necessary to expand the sensing frontier: thanks to data mining techniques e-mail, phone activity, or browsing behaviour can turn personal devices into mood detectors, cars can record illegal activity. 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, and so on) and expanding frontiers in datalogical sensing (the ability to extrapolate from the data provided by the existing technology). In this sense, the data mining process itself expands and fills in the available dimensions of sensing.

b) The Explosion of Sensor-Driven data

The logic of sensor related technology is circular and continuous – more sensors create more data which in turn open more avenues for new data collection by newly developed sensors – and the possibility of new correlations (thus new frontiers in "datalogical" sensing). The underlying logic of sensor permeation in the sensor society therefore requires the evolutionary development of sensors and sensor driven data alongside that of analytical techniques. This evolving environment highlights the increasingly pressing problem posed by the proliferation and permeation of sensors, data collection and data storage technologies enabled by the development of networked digital technology. In essence, the amount of data collected on a daily basis is now historically unprecedented, but is nonetheless, a small foretaste of things to come in the sensor society.

IBM claims, for example, that every day about 2.5 quintillion bytes of data is generated -- the data equivalent of a quarter million Libraries of Congress (IBM, 2013) and 90 percent of the world's stored data has been created in the past 10 years (IBM, 2013). That is, if human history were shrunk down to the length of a day, the vast majority of its accessible stored data would have been created in the equivalent of the last few seconds. Facebook alone reportedly enters the equivalent of about 50 Libraries of Congress into its databases each day (Kern, 2012). Historically, of the content of recorded data featured detailed records, books, and other storage media that took time and effort to create. Now databases continually fill up with data that is generated mechanically and automatically by a burgeoning array of digital sensors. Sensor driven data accumulates faster than human hands can collect it and faster than human minds can comprehend it.

This rapid redoubling of the world in informational form is the second attribute of the "sensor society." Bill Gates gestured in this direction when he described his version of a fully "documented life" in *The Road Ahead* (1996, 303). Gates envisioned a world in which all our actions, movements, interactions, communications, vital signs, and so on are automatically captured and stored so that they can be recalled at will to reconstruct the rhythms and events of our lives. In Gates's version of the monitored life, individuals would have control over their documentation and the recording devices that generated it: our information would be housed in our devices so that we could access it. Data generation in the sensor society pushes in the direction of fully documented lives and, further toward a fully documented world. However, neither the infrastructure nor the data is fully accessible or comprehensible to the individuals whose information is collected.

The sensor society records details about the information capture and recall process itself, registering the fact that a piece of data has been stored and tagged. Consequently, the sensor society remembers how it remembers which again leads to the combined proliferation of new sensors and data. This in turn fuels a tendency towards the self-generating automation of the self-remembering processes of sensor-driven data collection, information analysis, and predictive response.

The continual and self-generating machine-processing of data gives 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 metadata, 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 emails and online behavior for potentially useful patterns. The implicit logic of Google's rejoinder is that people should not care about having their information processed because no one is reading and comprehending the fully articulated content of their communications. 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, and so on).

We contend that it is precisely the potential of 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 email as metadata 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's second attribute is that content can be treated as metadata, 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).

We argue that Pentland's distinction does not quite 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 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" (2013, 298) – what Pentland describes as "behavior." As Packer 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" (2013, 298).

This is what Pentland most likely means when he says that Facebook posts and search requests are not of interest: that 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) as they are divorced from the ideational content of the message. Similar to Google's defence, Pentland's contention is representative of the process of meta-datafication.

There is a significant body of developing work that reveals the potential power of metadata to reveal all kinds of detailed personal information about individuals. Metadata can lead to the very content from which they allegedly distinguishes themselves. The notion that metadata is somehow less "private" than the content with which they are associated has come under considerable scrutiny (Ohm, 2010, Narayanan and Shmatlkov, 2010) Former Sun Microsystems engineer Susan Landau, for example, confided to the *New Yorker* magazine that the public "doesn't understand," that metadata is "much more intrusive than content" (Mayer, 2013). It is possible to unearth quite intimate details about individuals without having a human actually read their communications.

It should come as no surprise that, from a privacy perspective, the process of meta-datafiction erodes the concept of information privacy and the laws that flow from the concept. At the heart of information privacy law is personal information, otherwise known as personal data or "personally identifiable information." If information is not classed as personal information, information privacy law will not apply. (Schwartz and Solove, 2011) What is or is not personal information is therefore a threshold question which explains the importance of meta-datafication for its proponents. If ideational or content can be reconstituted as meta-data then potentially information privacy laws may not apply.

For example, 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 is consequently information about individuals or information that relates to individuals (refs). Personal information can therefore be specific or combinations of data that can identify individuals directly such as full name, drivers licence or social security number but can also be data that indirectly identifies individuals. For example, a residential address can be used as an aggregation point to aggregate different sets of data to enable identity. The legal definitions of personal information recognise 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 (Nissenbaum, 2010).

The process of meta-datafication is therefore an essential process in the self-generating environment of the sensor society. It is essentially the defence and in-built justification for ever-expanding, self-generating data collection and storage processes. However, we contend the opposite. The expansive logic of the sensor society creates the prospect that individuals will be uniquely identifiable from the meta-data created by sensor devices and sensor networks. The explosive nature of sensor derived data means that patterns of movement or online search behavior, or even unique typing patterns will give rise to the identification of individuals, especially in an environment where more and more sensors collect more and more data. As data from different sensors is combined and mined, it is possible to infer information about individuals – including details that would, in other contexts, fall into protected categories -- without needing to know their names or their addresses. Given the ease with which this data can eventually be traced back to individuals by drawing upon expanding aggregations of data, all data about persons harvested from increasingly comprehensive sensor networks are likely to become, for all practical purposes, personally identifiable. The notion that all data have the capacity to be personal information has important consequences in relation to the third attribute, the reliance of and faith in processes of predictive analytics.

c) The search for unintuitable correlations

The first two attributes of the sensor society contend that the proliferation of sensors causes an explosion of sensor-driven data. Attribute three focuses on the process of analysis, the process of making sense of sensor derived data. That process, we contend, is predicated on the rationale of predictive analytics and the constant search for unintuitive correlations. In other words, attribute three re-examines the very basis and justification for 'Big Data'.

New forms of data mining mean, "moving away from search as a paradigm to pre-correlating data in advance to tell us what's going on" (Hunt, 2012). In the context of policing and security, for example, data mining, or "pre-correlating data," as Gus Hunt, the CIA's Chief Technology Officer puts it, relies on the collection of as much data as possible before particular suspects or risks are known and then using this information to help predict possible suspects and threats. *All* data is 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, 2012).

This version of "total information awareness" does not really mean that everyone who is monitored is a suspect so much as it suggests that information about non-suspects and suspects alike (that is, everyone) is needed to help distinguish between the two. The functionality of attribute three is dependent upon access to and acquisition of all data which again re-emphasises the need for meta-datafication and attempts to avoid information privacy law, as highlighted in the previous section.

Along with information privacy law implications, the characteristic challenge 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 US 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 alia, 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 24-hour, multi-camera video surveillance. As one news account put 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-WWII 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. Semi-automated and automated forms of data sorting have, up until relatively recently, relied largely upon so-called "search and winnow" strategies (ref): sifting through information troves using keywords to isolate findings that might be of interest to analysts.

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. There is an element of randomness in this speculative rationale: a particular piece of information might eventually prove useful once other, yet-to-be collected data sets are amassed; but it might not (and there is no way to know in advance). Nevertheless, 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. Moreover, there is no way to definitively rule out the possibility that new data might make currently useless data useful -- hence the goal of holding on to data "forever." Even when data is intentionally collected, the specific reason for its collection – it's potential usefulness -- can remain deferred, perhaps indefinitely.

By the same token, the data mining processes that have developed to handle the large amounts of data generated by sensors are *emergent* ones because their goal is to generate un-anticipatable and un-intuitable correlation: that is, patterns that cannot be predicted in advance. The analytics therefore derive sense of the explosion of data generated by the proliferation of sensors. That said, the analytics are also a key component in the quest for *all* data as un-anticipated or un-intuitive results can only be derived from new data, even seemingly irrelevant data. For example, in Section 2 we highlighted the company Evolv, that mines large sets of recruitment and workplace data, including the browser used to submit job applications. The predication that certain browser usage indicates certain employee traits is potentially useful for employers interested in screening job applicants, but it is not the result of a strategic and systematic deployment of monitoring capabilities. Instead, it is the serendipitous result of data that was collected as a by-product of the job application process.

In the sensor society, internet browsers (in conjunction with the computers that run them and the networks that carry their data) therefore serve as sensor networks, insofar as they can relay information about users back to (in the Evolv case) prospective employers. 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. The finding, in other words, was the result of a fishing expedition based on trolling through some 3 million pieces of data about 30,000 employees (*The Economist*, 2013). 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 generated by the available sensing infrastructure. 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. In other words, the collection of browser information is part of the quest to accumulate all data, from all sensors to fulfil the unquenchable purpose of predictive analytics.

Attribute three of the "sensor society" thus refers not just to the proliferation of automated sensing devices across the landscape, but also to the associated logics that are characteristic of automated, mechanized sensing: always-on information capture, the associated avalanche of data, and the consequent tendency toward automated information processing and response. It also refers to the changing ways of thinking about and using information with which automated sensing is becoming associated: in particular the Janus-faced use of data for both prediction

and information retrieval – that is, the related goals of modelling the future and reconstructing the past. In order to fulfil both of these requirements, sense-making infrastructures are needed, which brings us to an additional significant attribute of the sensor society.

d) Sense-making infrastructures

As the data from sensors accumulate, they can be used not just to model the future, but to mine the past. Consider, for example, the use of mobile phone 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). For the first three attributes of the sensor society to function effectively, a complete archive is required to supplement the vagaries of reported actions and memories by externalizing them in the form of machine readable databases.

The issue of infrastructure is accordingly central to these examples and thus the sensor society. Infrastructure in this sense enables the first three attributes of the sensor society by facilitating the reconstruction of the past and predictions of the future. It is the infrastructures of sensors that generate sensor driven data. The infrastructures of collection enable the explosion of collectable data. The infrastructures of prediction enable the sense-making of sensor derived data and thus give purpose to sensors.

Jeremy Packer 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" (2013, 297).

In the case of the infrastructural and retroactive function of the sensor society this notion of *reassemblage* refers literally to the piecing together of data to make sense of the past and to predict the future.

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 characterize what might be described as the 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 alia, 1996). Even Hardt and Negri's 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" (2009, 294). 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 sensor society, by contrast, redirects our attention to infrastructures that make data collection capture, storage, and processing possible, and the relations of ownership and control that shape who has access to data and who sets the parameters and priorities for its use. 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 figures 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). 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). In other words, the sensor is inextricably related to the data, the analytics and the infrastructure otherwise the sensor is merely a non-sensing device. Likewise, the infrastructure and the analytics become the justification for the collection of data derived by the sensor.

Sensors can of course operate distinctly at close range. A device, for example, can determine to brighten the screen on a smart phone in sunny conditions, or the fingerprint reader that unlocks a laptop. However, it is when these sensors become fused with data generation, analysis and infrastructure that the salient characteristics of the sensor society emerge. In keeping with our goal of proposing fruitful angles for approaching issues of monitoring, surveillance, privacy, and control, the final substantive section explores how the four attributes of the sensor society combine to provide a insight into recent technological developments.

4. Making Sense of the Sensor Society

The real value of the sensor society, as we have already suggested, derives from viewing the four attributes together. 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 different but inter-connected infrastructures. The logic of targeting in the sensor society shifts: individuals are not singled out for prescribed monitoring for a specified purpose. Rather they 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 forms in conjunction with that of others. In the sensor society, the target is the pattern and the pattern is always emergent. Hence the need for accumulation and aggregation of not just all existing data but all new data as any new piece of data will help to generate a new target pattern and thus fulfil the purpose of the infrastructure's goal.

Conventional understandings of privacy as control over one's self-disclosure and self-presentation are complicated by this reconfiguration of targeting towards 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 logically follows that data collection and analytical infrastructures are equally opaque. The legal theorist Tal Zarsky (2013) 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)" (1519).

As such, processes of opacity that yield unanticipated uses for data which result in uninterpretable decisions undermine some of the key foundations of information privacy law, namely, informed consent and even ideas such as contextual integrity (Nissenbaum, 2010). Moreover, as highlighted in our discussion about sensor derived data and predictive analytics, all data needs to be treated as personal data in the sensor society as any given piece of data, aggregated with other pieces of data for the purpose of predictive pattern generation, could have the capacity of identifying an individual but more importantly, could be used in a way that detrimentally impacts on their life. The multiple variables that Zarsky rightly highlights are not pieces of abstract data. They are representations of our life which can be used to our benefit and detriment in unintuitable and unknown ways. For example, when one prospective employee does not get the job they desire due to the browser they use.

Neither the concept of information privacy law nor anti-discrimination law are designed to cope with the vastness of data collection and analysis presaged 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 is therefore extremely complex due to the non-targeted nature of actions and intentions in legal frameworks largely predicated on liberal biases that seeks to govern and protect intrusions against individuals (Cohen, 2012).

The problematic nature of specified regulatory intervention leads to the eventual overspill of sensorised activities, logics and infrastructures which explains the current proliferation and creep towards the sensor society. As Packer (2013) puts it, "The breakthrough of digital media…all of reality -- is now translatable. The world is being turned into digital data and thus transformable via digital manipulation" (298). The result is that so-called "function creep" is not ancillary to the data collection process, but is built into it: the function *is* the creep. This fact is not simply the result of sensor driven generated data which facilitates the concomitant processes of data storage, sharing, and processing. Rather, it is the fact that search for unintuitive correlations, and the infrastructures that facilitate predictive analysis are required to view every piece of data as useful and essential for the generation of an un-anticipatable pattern.

It is our contention 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 generation, data collection, predictive analysis, and response. The sensor society is consequently infrastructural in nature. It is in the infrastructures underpinning the sensor society where the ever-expanding, self-generating true complexity of the society unfolds. These infrastructures give sense to sensors and sensors give justification for infrastructures. The sensor society looks beyond the ephemeral construct of 'Big Data' and leads us to critically question the power structures behind infrastructural creation, development and implementation. Ultimately, therefore, the sensor society is about a complex dispersion of power inherent in the newly developing sensorised processes of life.

To propose 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 of 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 positing the notion of a sensor society we have highlighted important issues facing those interested in topics 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 develop their theoretical and conceptual imagination to keep pace with the technology and its deployment.

Works Cited:

Beniger, J. R. (1986). *The control revolution: Technological and economic origins of the information society*. Harvard University Press.

Biosenics (2014). http://www.biosensics.com/

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/.

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.

Chakrabarti, S. (2009). *Data mining: know it all*. Morgan Kaufmann.

Clarke, Roger (2003) Dataveillance --15 years on. Accessed at: http://www.rogerclarke.com/DV/DVNZ03.html.

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. 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.

Daley, Jason, Adam Piore, Preston Lerner, Elizabeth Svoboda (2011). "How to Fix Our Most Vexing Problems, From Mosquitoes to Potholes to Missing Corpses."

Discover, October. At: http://discovermagazine.com/2011/oct/21-how-to-fix-problems-mosquitoes-potholes-corpses#.UqlKfmQW3EU.

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 September 2, 2013).

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

Drake, Thomas (2013). "Snowden saw what I saw: surveillance criminally subverting the Constitution." *The Guardian*, June 12, retrieved online at:

http://www.theguardian.com/commentisfree/2013/jun/12/snowden-surveillance-subverting-constitution.

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, p. B1.

Dyson, E., Gilder, G., Keyworth G., & Toffler, A. (1996), 'Cyberspace and the American Dream: A Magna Carta for the Knowledge Age', *The Information Society*, 12:3, pp. 295-308.

Economist, The (2013). Robot recruiters: How software helps firms hire workers more efficiently, April 6. Accessed online at: http://www.economist.com/news/business/21575820-how-software-helps-firms-hire-workers-more-efficiently-robot-recruiters.

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

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.

Gates, Bill (1996). *The Road Ahead* (New York: Penguin).

Hardt, M., & Negri, A. (2009). *Empire*. Harvard University Press.

Hotz, Rober Lee (2011) The Really Smart Phone. The Wall Street Journal, April 23, http://online.wsj.com/news/articles/SB10001424052748704547604576263261679848814.

Hunt, Gus (2012). "Big Data: Operational Excellence Ahead in the Cloud," Presentation to the Amazon Web Services Government Summit 2011, 26 October, Washington, D.C., http://www.youtube.com/watch?v=SkIhHnoPpjA (accessed 10 August, 2012).

IBM (2013). The IBM Big Data Platform, IBM Software Group. Retrieved online at: http://public.dhe.ibm.com/common/ssi/ecm/en/imb14135usen/IMB14135USEN.PDF.

Kern, Eliza (2012). "Facebook is collecting your data – 500 terabytes a day." *GigaOm*, August, 22. Retrieved online at: http://gigaom.com/2012/08/22/facebook-is-collecting-your-data-500-terabytes-a-day/.

Kolakowski, Nick (2014). "South Koreans Using Kinect to Monitor DMZ". Slashdot, 3 February, http://slashdot.org/topic/cloud/south-koreans-using-kinect-to-monitor-dmz/

LiKamWa, R. (2012). *MoodScope: Building a Mood Sensor from Smartphone Usage Patterns* (Doctoral dissertation, RICE UNIVERSITY).

Lyon, D. (2001). *Surveillance society*. Buckingham: Open University Press.

Lyon, D. (2008). "Surveillance Society". Talk for Festival del Diritto, Piacenza, Italy.

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, 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.

MIT Media Laboratory (2011). Sociometric Badges. Web site, accessed at: http://hd.media.mit.edu/badges/.

Narayanan, A and Shmatlkov, V (2010). Myths and Fallacies of Personally Identifiable Information. Communications of the ACM. http://www.cs.utexas.edu/users/shmat/shmat_cacm10.pdf

Nissenbaum, H. (2010). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford, Calif.: Stanford Law Books.

Ohm, Paul (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. UCLA Law Review 57, 1701.

Packer, J. (2013). Epistemology Not Ideology OR Why We Need New Germans. *Communication and Critical/Cultural Studies*, *10*(2-3), 295-300.

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).

Schermer, Bart (2008). "Privacy and Visibility in the Sensor Society" http://www.slideshare.net/Considerati/privacy-and-visibility-in-the-sensor-society

Schwartz, Paul and Solove, Daniel (2011). The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review* (86), 1814.

Sledge, Matt (2012). "CIA's Gus Hunt on Big Data." *Huffington Post*, March 21. Accessed online at: http://www.huffingtonpost.com/2013/03/20/cia-gus-hunt-big-data_n_2917842.html.

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.

Waber, Ben (2013) People Analytics. FT Press.

Webster, F. (2007). *Theories of the information society*. Routledge.

Williams, Bryan Glynn (2013). "Inside the Murky World of 'Signature Strikes' and the Killing of Americans With Drones," *The Huffington Post*, May 31, http://www.huffingtonpost.com/brian-glyn-williams/nside-the-murky-world-of-_b_3367780.html (accessed September 2, 2013).

Wood, D. M., Ball, K., Lyon, D., Norris, C., & Raab, C. (2006). A report on the surveillance society. *Surveillance Studies Network, UK*.

Zarsky, Tal (2013). "Transparent Predictions." *University of Illinois Law Review*, 4, 1503-1570.