Article

International Federation of Library Associations and Institutions
2018, Vol. 44(3) 229–239

DOI: 10.1177/0340035218778054
journals.sagepub.com/home/ifl

Privacy, obfuscation, and propertization

Tony Doyle
Hunter College, New York, USA

Abstract

As our digital wake ripples out, big data is standing by to ride it, applying its analytics to make unnerving inferences about our characters, preferences, and future behavior. This paper addresses the challenge that big data presents to privacy. I examine what are perhaps the two most promising attempts to repel big data's attack on privacy: obfuscation and the "propertization" of personal information. Obfuscation attempts to throw data collectors off our digital trail by confusing or misleading them. Propertization calls for treating personal information as intellectual property and would require that data holders compensate data subjects for any secondary use. I try to show that both defenses largely fail. I conclude that privacy is a lost cause and that we should call off the attempts to defend it from the moral point of view. I close with some thoughts about what this all means for libraries.

Keywords

Autonomy, big data, information technology, obfuscation, privacy, propertization

Submitted: 27 December 2017; Accepted: 25 April 2018.

Introduction

Big data, with its massive collection, thorough aggregation, predictive analysis, and lightning dissemination of personal information, has produced previously unfathomable benefits and insights. It is fomenting a Copernican revolution in the social sciences (Mayer-Schönberger and Cukier, 2013; Stephens-Davidowitz, 2017). It has also been a boon for detecting credit card fraud and money laundering, monitoring traffic flows, refining digital translation, matching consumers with useful products and services, improving diagnoses, and tracking public health trends (Acquisti, 2014; Barocas and Nissenbaum, 2014a; Mayer-Schönberger and Cukier, 2013; Schneier, 2015). In libraries, the user experience can be enhanced through "personalization," in which items are recommended based on a user's previous interests, on what other users with similar interests have sought, or on friends' preferences (Pekala, 2017). Analysis is replacing intuition; the gut is yielding to algorithm.

But big data is bearing down on privacy. In what follows, I will use the expression *big data* to cover not just the collection but also the analysis of data. After all, it is the monumental harvesting in combination with ever more sophisticated analysis of it that poses the real threat to privacy.[1] Once upon a time there was too much data to save. No more: Storage costs have long been in free fall. These days, even if a stockpile's utility is not immediately apparent, data holders will warehouse it in the reasonable hope that uses will emerge (Angwin, 2014; Mayer-Schönberger and Cukier, 2013; Schneier, 2015). Gone, too, is the default protection of file cabinets, paper archives, and stand-alone computers. Today, much of that information is available from a single point. Once aggregated, the trove can yield novel and uncanny inferences about our activities, preferences, commitments, aspirations, vulnerabilities, and future behavior. This bounty can then travel widely in a flash. If not quite an open book yet, our lives are anything but a locked diary.

Corresponding author:
Tony Doyle, Hunter College Philosophy Department, Hunter College Library, 695 Park Avenue New York 10065, NY, USA.
Email: tdoyle@hunter.cuny.edu

How should we respond? Notice and consent (or choice), once the great market-based hope for privacy protection, has failed signally. The idea is that optimal privacy is reached when informed individuals agree, or not, to the collection and subsequent use of their data (Schwartz, 2004). However, notices that attempt to cover all contingencies are unreadably long and would gravel even experts. For instance, a 2008 study showed that the average American would need 244 hours to plow through the privacy policies of all the websites he accessed in a year, at an opportunity cost nationwide of $781bn (cited in Landau, 2015). Pared down, accessible statements, on the other hand, strip away crucial detail (Nissenbaum, 2011). Either way, refusing to accept the terms means that we cannot use the service. Third, much of the value of information comes from secondary uses that not even data scientists can anticipate at the time of collection, making nonsense of consent even when people comprehend the statements (Barocas and Nissenbaum, 2014a; Meyer-Schönberger and Cukier, 2013). Fourth, by agreeing to share information, say, about my health, brokers can proceed to make non-trivial inferences about the health of others not in the data set but who are otherwise akin in age and habits (Acquisti et al., 2016; Barocas and Nissenbaum, 2014b). I say more about this below.

Legislation restricting what data holders can do with personal information might offer some hope. However, as Finn Brunton and Helen Nissenbaum point out (2011 and 2015), the law will inevitably lag behind the breakneck innovations of privacy-threatening technology, and data brokers and their clients are likely to have undue influence on how legislation is crafted and how vigorously it is enforced. Inadequate, too, are measures like the US Federal Trade Commission's Fair Information Practices (FIPs) of 1973. These principles forbid secret record-keeping and the unauthorized secondary uses of personal information, both of which are routinely violated. They also give data subjects the right to correct any errors in their dossiers. However, if people do not know what records about them exist, this protection comes to nothing. Anyway, errors are not the main problem; harmful inferences from accurate information are. In addition, the FIPs propose to make data holders responsible for any harm resulting from misuse of data. But the harms in question, which I discuss below, are generally cryptic and amorphous, thwarting detection and frustrating enforcement.

That leaves two far more plausible responses to the assault of big data: obfuscation and propertization. Obfuscation attempts to shield privacy by producing plausible but "misleading, false, or ambiguous data" about a person, "with the intention of confusing an adversary or simply adding to the time or cost of separating good data from bad" (Brunton and Nissenbaum, 2011). Propertization proposes treating personal information as a kind of intellectual property and then compensating data subjects for its use. Ideally, obfuscation and propertization both will allow the rest of us to determine how much privacy we want to retain. I deal with each in turn, concluding that neither will seriously shelter privacy from big data's plunder. First, though, I would like to say a bit more about privacy and about how big data threatens it.

Privacy

I will pass on defining *privacy*, since any definition that I offer is bound to be open to plausible counter-examples. For the sake of discussion, I will follow Helen Nissenbaum's (2010) focus on the context-bound, morally permissible flow of personal information, which she calls "contextual integrity." I will accept her suggestion that sound privacy protection is a function of whether the flow of information follows applicable, context-bound norms or "transmission principles." Each social context comes with its own set of norms. Thus, the norms governing, or constraining, the flow of information differ according to whether the context is, say, doctor-patient, teacher-student, employer-employee, or friend-friend. Transmission principles include *consent*, *confidentiality, reciprocity, notice,* and *desert*. For any bit of personal information that passes hands we need to ask the following questions:

- What is the information about?
- Who is it about?
- Who receives it?
- Under what circumstances?

When the appropriate transmission principles are adhered to, contextual integrity is preserved, otherwise not. Violations of contextual integrity are prima facie wrong; that is, they alert us that the flow of information in question raises serious, though not conclusive, moral reasons for not engaging in the practice. For instance, confidentiality in the context of healthcare means that my doctor is forbidden from sharing information about my health with my employer or *The New York Times* but is free to pass it along to my insurance company or appropriate specialists (Samuelson, 2000). Reciprocity reigns with close friends but not with one's doctor. If I ask a good friend about his health, it can be appropriate for him to ask me about mine in response. By contrast, although my doctor deserves an honest account

in response to her questions about my health, it would obviously be out of place for me to inquire about hers (Nissenbaum, 2010)!

Nissenbaum proposes contextual integrity in response to the radical alterations in the flow of information that digital technology has wrought. Specifically, problems arise when information collected for one purpose—for example, from an app that monitors exercise patterns—is used for other purposes, like determining car insurance rates or making hiring decisions. Contextual integrity implies that it is impossible to say beforehand whether a given piece of information is private or sensitive, on the one hand, or public or non-sensitive, on the other. Context gets the last word. New Year's Eve snapshots of me with lampshade may reasonably be shared among my family and closest friends but not with my boss or my students. Similarly, no information is inherently public. It all depends on who gets it and how they handle it once they do. That is, it is strictly a matter of the context in which the information flows. That I invariably buy black T-shirts and whitening formula toothpaste might seem non-sensitive, but the pattern might slot me with reckless drivers and boost my premiums, despite my own perfect driving record. My stroll down Main Street yesterday might seem public if anything does. After all, it was broad day, hundreds saw me, and I wore no disguise. However, the presence of face-recognizing surveillance cameras there can still violate contextual integrity. First, I might not even be aware of the cameras (*notice*) or, if I am, know nothing about the fact that the information gleaned can be combined with still more information about me and relevant others and then widely disseminated (*notice* again). Second, even if I am wise to all this, I am not given the chance to say yes or no to the capture and subsequent use of the information so gathered (*consent*). Big data menaces privacy because it regularly transgresses time-honored constraints regarding who should get certain information, under which circumstances, and what they can do with this information once they have it. In fact, Nissenbaum's theory might almost be called *beyond privacy,* given her emphasis on morally appropriate information flow. Consider the fact that people who own Harley-Davidson motorcycles tend to have lower than average IQs (Stephens-Davidowitz, 2017). Prospective employers could use that information to exclude qualified candidates from the interview pool. It is not clear that the victims' privacy has been violated. Still, contextual integrity has been breached, since information gathered for one purpose has been used, without notice or consent, for another purpose, to the data subjects' detriment. Anyway, if we look at enough characteristics, we are bound to find some that happen to correlate with traits like IQ (Stephens-Davidowitz, 2017).

The foregoing implies that privacy is a normative concept, which raises the question, Why value it? My answer is that it tends to promote autonomy (see Cohen, 2000). Autonomy means being able to make choices, free of coercion or manipulation, in the light of one's own considered conception of the good life. Autonomy, for its part, promotes well-being by enabling us to increase our opportunities and advance our projects (Tavani and Moor, 2001). Commercial tracking, monitoring, and profiling are bad insofar as they tend to be inimical to privacy and thus to autonomy. People are generally better off when they have more rather than less of both. When information technology threatens them, general well-being is undermined. Privacy matters.

Big data's revelations: Further examples

Big data's phenomenal success comes from taking piles of data collected for one purpose, for example the location information needed to route your calls or texts, and applying them to myriad, apparently unrelated, secondary purposes, like predicting where you will likely be next week at this time. Big data's trick is to merge discrete and apparently trivial details from a person's life into a coherent and potentially privacy-threatening whole that is greater than the sum of its parts (Mayer-Schönberger and Cukier, 2013; Nissenbaum, 2010). This process enables data holders to discriminate ever more finely among people to arrive at the optimal decision, from the former's point of view, about how to treat you and me at a given time (Rule, 2007). For instance, since the early 1990s insurers have used credit scores to figure out who to write policies for and what to charge for the policies they do write, since people with bad credit are significantly more likely to make claims than those with good (Rule, 2007). More recently, data miners have honed their technique to reveal, for instance, that folks who buy cheap motor oil, Chrome-Skull car accessories, and hang out in the local bar, tend to have bad credit and presumably are bad insurance risks as well. This cohort's mirror image are those paragons who buy home carbon monoxide sensors, snow roof rakes, felt feet for their furniture, and premium bird seed (Duhigg, 2009; Mayer-Schönberger and Cukier, 2013).

That's not all. As Cathy O'Neil (2016) amply documents in *Weapons of Math Destruction*, it turns out that just about any data is credit data. In the United States, it is illegal to use credit scores without

subjects' consent. Lacking legal access to the reports themselves, data handlers have helped themselves to . . . you name it: zip codes, purchases, places shopped, Internet surfing patterns, or having friends—real or social media—who meet certain criteria. With this information, data handlers can concoct an e-score, a data-rich stand-in for creditworthiness (Mayer-Schönberger and Cukier, 2013; O'Neil, 2016). E-scores enable data holders to sidestep consent for access to credit scores. Creditworthiness, or its e-score facsimile, in turn substitutes for other virtues like trustworthiness and dependability on the one hand, or for a multitude of sins on the other, whether one is guilty of them or not. One's creditworthiness, real or apparent, can then be used to determine whether one gets a job, a loan, an apartment, or, of course, insurance (and at what rate). In some parts of the United States, creditworthiness counts for considerably more than driving record in determining car insurance rates. O'Neil (2016: 165) adduces Florida, where in 2014 "adults with clean driving records and poor credit scores paid an average of $1,552 more than the same drivers with excellent credit and a *drunk driving conviction*." All of this proceeds, in most of the United States at least, with near impunity, despite the fact that one's credit rating or e-score can slip or tumble for all kinds of reasons that have nothing to do with bad behavior or a weak character, like a devastating accident or serious illness.

The apparently innocuous data that we generate as we go through the motions is more or less up for grabs, and in critical mass it enables data holders to categorize us according to race, ethnicity, political views, and sexual orientation, as well as according to more specific criteria like *gambler*, *smoker in the house*, *adult with elderly parents*, and *adult with wealthy parents* (Singer, 2013). The categorization affects the ads or job offers we see online, the products and prices we are offered there, and the quality of service we receive in a call center (Angwin et al., 2017; O'Neil, 2016).

This is the panoptic sort that Oscar Gandy (1993) warned of long ago. The techniques of big data permit the classification of people based on "their estimated presumed economic or political value" (p. 1). Big data's ability to do so has improved dramatically since Gandy wrote, thanks to those plummeting storage costs, greatly expanded networks, and ever-more sophisticated techniques of re-jiggering data, from which precise, surprising, and profitable inferences can be made about you and me. Gandy (1993) calls the panoptic sort a "difference machine," a "discriminatory technology," that "allocates options and opportunities" based on personal characteristics (pp. 15 and 17). The sort is "an integrated system that is involved in the identification, classification, assessment, and distribution of individuals to their places in the array of life chances" (Gandy, 1993: 35). In other words, it has a great deal to say about the odds that a person has of living a good life. The terms of the exchange are set by data holders and their clients. Nearly all of this happens without data subjects' consent or even awareness of what is collected, who it is being shared with, or what those third parties are doing with the information once they have it (Gandy, 1993). They can see us, but we can't see them. The new panopticon makes Bentham's prototype seem quaint.

Again, big data is all about effective discrimination: Businesses quite reasonably want to know both who to seek out and who to avoid. The reward for effective discrimination is increased profit (Rule, 2007; Schneier, 2015). As we just saw, the canny third party need not have any information about our actual characteristics. Information about those who are otherwise like us suffices to sort us in all kinds of ways. For instance, frequenters of gambling sites might be a bad risk for a bank loan (Steel and Angwin, 2010). More subtly, a detailed picture of one's health can emerge without any third party access to one's medical records, dodging consent. Obesity, a handy proxy for a suite of health risks, can be reliably inferred from the following: regular fast food dining, frequent online shopping for clothes, being a childless minivan owner, and subscribing to premium cable (Walker, 2013). One data broker was able to identify people who were probably arthritic by looking at cat ownership, a preference for jazz, and participation in sweepstakes (Walker, 2013). Creditworthiness, exercise habits, recent websites visited, and TV watching habits might even be interchangeable in some cases with blood and urine samples as a predictor for heart disease, high blood pressure, diabetes, or depression (Mayer-Schönberger and Cukier, 2013; Schneier, 2015). The same goes for race: Zip code coupled with mother's level of education say a lot about it (Ohm, 2014). So much for consent and the control over personal information that it was supposed to provide. In fact, the production of most new information about me can now proceed without my consent and even without information about the relevant trait at all (Mai, 2016).

Obfuscation

Finn Brunton and Helen Nissenbaum (2011, 2013, 2015) offer obfuscation as a rejoinder to big data's outrages. Obfuscation makes the collection of data

about individuals "more difficult to act on, and therefore less valuable . . . adding to the cost, trouble, and difficulty of doing the looking" (Brunton and Nissenbaum, 2015: 46–47).

The case for obfuscation

Online obfuscation promotes anonymity or hides one's actual searches among a gang of plausible fakes. In short, obfuscation makes it harder for receivers of information to tell signal from noise, wheat from chaff (Brunton and Nissenbaum, 2013). It is a "troublemaking strategy" that lets those who use it buy time or wall themselves off from importunate or malign third parties (Brunton and Nissenbaum, 2015: 4). Think of drawing the shades, donning a disguise, or chatting in a language that most others do not understand. It might be our best hope for keeping our information from the clutches of big data.

The technique is actually as old as the hills. Natural selection has gone in for it time and again. Consider the monarch and viceroy butterflies. As a result of feeding on milkweed as larva, monarchs are toxic to many vertebrates (Oberhauser, 2011). The species advertises its venom in flashy black and orange. A bird that has tried to snack on a monarch in the past will presumably remember the shock and shun similarly colored butterflies in the future. It is even possible that natural selection favors predators that are averse to eating monarchs, or anything like them, from the start. At least one mimic has capitalized on the monarch's combination of showiness and bad taste: viceroys (Schnur, 2002). The non-toxic viceroys are all but indistinguishable from their noxious cousins. It is easy to see why natural selection might incline towards obfuscation here. For the predator, information about potential quarry is ambiguous. Is the vibrantly colored bug up ahead a hearty lunch or a possible last meal? The savvy hunter will avoid anything turned out like a monarch.

Online obfuscation works similarly, attempting, for instance, to cloak the surfer's identity or the nature of her queries enough to throw unbidden third parties off the trail. The point is to drown the signal out with ever more noise (Brunton and Nissenbaum, 2011; Howe and Nissenbaum, 2009). Take the web-based obfuscator, Tor. Once I join the Tor network and allow my machine or device to function as a relay, my queries are encrypted and are received not from my IP address but from another "node" in the Tor relay network. The response comes back to me via other nodes, thereby shrouding my identity. Not only can snoops not decrypt the message, but because my computer is acting as a relay, they also will not know whence it came. As Brunton and Nissenbaum (2015: 20) put it, my messages are now "safe in a flock of other messages" that I and others in the network pass along. The result is that adversaries are far less likely to tie my web activity back to me than they would be without the obfuscation. If it is all right to disguise my appearance in public, particularly in the light of proliferating video surveillance, then it looks like I am justified in obfuscating my online activities, or even concealing my identity there altogether, to dodge monitoring and profiling. Until data collectors or regulators can guarantee that personal information flows within the bounds of contextual integrity, those concerned about their privacy apparently have little choice but to obfuscate.

Problems for obfuscation

Nevertheless, obfuscation faces challenges. The first is moral and has to do with the free ride that obfuscators seem to enjoy. The Internet is for the most part ostensibly free because the vast majority of people either innocently share or are coerced into parting with reams of information when they go online. Obfuscators, unlike their credulous or uninformed counterparts, get all or most of the benefits of the Internet without paying for them in the coin of surveillance. Take ad blockers. The software hides ads from the user, while clicking on them all, thereby obfuscating users' true interests and preferences. The result is an ad-free Internet experience. A similar point could be made about Tor: A user cannot be targeted by the personalized ads that underwrite a "free" Internet if she or her true activity is invisible to marketers. It looks like obfuscation offers a haven to its practitioners while abandoning everyone else to the choppy sea. Does this mean that obfuscators are "sneaks more than rebels," as some critics have suggested? (Brunton and Nissenbaum, 2015: 67; see also Brunton and Nissenbaum, 2013). Not necessarily. True, Brunton and Nissenbaum concede, the free Internet is sustained by user information. However, the terms of exchange between data subjects and data gatherers are baffling to most and invariably set by the latter. Normally, when we buy a product, we can form a pretty good idea of its value before paying up. This is not the case when we go online, where hidden costs abound: Most of us do not have the foggiest about how the capture and shuffling of our information will redound to us. Privacy partisans are not asking to get something for nothing. They can acknowledge that Google and Facebook will not provide their services for free and that the substratum of infrastructure involved in GPS services or connecting us to online

friends demand huge investment.[2] What they challenge is the price. As Brunton and Nissenbaum point out, when we trade information for a service or product, we are in effect handing a blank check over to data collectors (Brunton and Nissenbaum, 2015; Mayer-Schönberger and Cukier, 2013). Moreover, the price we are offered for products or services online can be adjusted according to algorithmic hunches about what we are willing to pay. The discrimination is opaque to consumers and perhaps even to merchants. It represents yet another blow for consent, since what I am charged might be a function of what others like me have been willing to pay in the past (Acquisti et al., 2016). Also, as Brunton and Nissenbaum (2011) point out, everyone, not just the cognoscenti, can obfuscate. It is a tool that "aids the weak against the strong" (Brunton and Nissenbaum, 2011), that is, those who know that they will be tracked but lack the skills to take stronger measures to defend their privacy (Brunton and Nissenbaum, 2015). Obfuscation is a way of standing up to the "coercion, exploitation, or threat" of big data (Brunton and Nissenbaum, 2015: 64). It enables us to snatch back that blank check before it is cashed.

Brunton and Nissenbaum attempt to clear obfuscators of the charge of free riding by pointing out that they are "not actively attempting to keep others from enjoying the same benefit . . . [They] cannot be expected to imperil themselves solely because others are in peril; they cannot be morally obligated to starve simply because others are starving" (Brunton and Nissenbaum, 2013: 179). The point is that obfuscators are leaving non-obfuscators no worse off than they would have been without the obfuscation (Brunton and Nissenbaum, 2015).

Will this wash? If the harms of obfuscation fall only to data holders and their clients, no problem, since, as Brunton and Nissenbaum (2015) rightly point out, the information exchange takes place under the dual asymmetries of power and knowledge: The information is often squeezed from us as a condition of countless routine transactions, and we generally do not know what happens to it or how its subsequent use affects us. However, although obfuscation is freely or cheaply available to all, it seems odd to describe the people who will in fact obfuscate as weak. After all, they will likely on average be highly educated and reasonably well-to-do. Brunton and Nissenbaum's slogan seems to be "let the devil take the hindmost" when it comes to evading surveillance. The fact is, they do not know what the costs of obfuscation are to non-obfuscators. They do concede that obfuscation needs to be judged case-by-case (Brunton and Nissenbaum, 2015). Still, the suspicion remains that obfuscators can enjoy a truly free Internet only if others are foolish or naïve enough to surrender their own information.

Also, Brunton and Nissenbaum never satisfactorily face up to the potentially great social costs of obfuscation online, particularly its ability to conceal grave misdeeds on the Web. Consider, for instance, the Silk Road website, which flourished from 2011 until 2013 and provided a massive venue for the sale of arms, human organs, and all manner of recreational drugs (Bilton, 2017). The founder and his associates used Tor to muddy their communications and traded exclusively in all but untraceable bitcoin. I am not claiming that a case like this shows that effective online obfuscation is unconditionally wrong, still less that it should be illegal. However, defenders of the practice need to deal with hard cases like this. Brunton and Nissenbaum have not.

Finally, even if its defenders can plausibly address the moral trials that obfuscation faces, they still have to deal with a serious practical one. As Brunton and Nissenbaum describe it at least, the practice protects at best our online activity. It provides no shield for the myriad other digital records that we routinely deposit through toll passes, credit cards, or our phones. Even if my Internet activities were maximally obfuscated, third parties would still be getting loads of data about me. And the web habits of comparable non-obfuscators will still enable data holders to draw many damaging inferences about me. So it looks like obfuscation is both morally suspect and in practice not terribly effective. In the light of these deficiencies, I would like to consider another option for at least reducing the profitability of big data: propertization, that is, assigning property rights in personal information to data subjects.

Propertization

Propertizers plausibly point out that big data is reaping most of the benefits of collection and analysis while bearing few of the costs, specifically to privacy (Laudon, 1996). These costs are externalized—that is, borne by data subjects—in the same way that polluters externalized theirs in the days before emissions were regulated or taxed. Currently, those who profit from the collection, analysis, and dissemination of personal information have little incentive not to sweep up as much as they can and sell it to the highest bidder. The propertizer need not be opposed to data holders profiting from collection and analysis. After all, the money to finance the ostensibly free Web has to come from somewhere, and those who provide these services naturally need a financial motive to do so. Also, the propertizer will concede that data brokers add considerable

value to the data they amass. The propertizer's objections relate only to *the extent to which* data mongers are profiting at the expense of data subjects.

Actually, propertization is already here to some extent, as when insurers give us discounts for letting them document our driving habits, keep track of how much we exercise, or monitor our cholesterol. Other examples include promo codes and loyalty cards, both of which tie discounts to our identity. Even "free" apps, as well as Google and Facebook, implicitly also involve propertization, since users are in effect swapping information about themselves for services rendered. The current proposal would simply formalize this arrangement and urge that we get our due. The information is fundamentally ours. Let the law acknowledge this.

The case for propertization

I will assume that the best argument for propertizing personal information disclaims moral rights and instead appeals to the good results of granting data subjects legal rights in their information. The same could plausibly be said of any system of property. It is justified to the extent that it contributes to overall well-being, otherwise not. Propertizers in particular argue that data subjects should control the disposition of their personal information, just as they do their house or car (Litman, 2000). Imagine that your barber was profiting from your trimmed hair, say, by making fine wigs from it or selling it to third parties, who were analyzing its DNA or testing it for what it revealed about your lifestyle. You would probably want a say in the matter. Why not with regard to the information you typically surrender in the course of a day? Second, since no two people will value their own information in exactly the same way, a market in personal information would allow them to assign different values to the same type. As Lawrence Lessig (2002: 262) puts it, "I may be a freak about people knowing my birthday, and so would never 'sell' access to that fact for any price, but someone else may be willing to sell access in exchange for 100 frequent flyer miles." I could agree to allow myself to be targeted by data collectors and their clients; you might absolutely refuse to do so. Propertization then would satisfy the full range of privacy preferences, from indifference to obsession (Lessig, 2002; Samuelson, 2000). Like most other market exchanges, propertization seems to offer a positive sum game between buyer and seller. Plus, presumably competition among collectors would drive the price of personal information up, meaning that less of it would make the rounds. It would also allow data subjects to get the best price they can (Rule, 2004). What are we waiting for?

Problems for propertization

A central assumption of propertization is that data subjects can give informed consent to the use of their information for a certain price. Informed consent in turn assumes that subjects are able to form a reasonable idea of what their information is worth at the time of sale, as is generally the case with, say, cars or houses. However, we have good reason to believe that the market will consistently undervalue personal information. At the very least, most sellers will not be in a position to know anything like what its true market value is. As we have seen, the value of personal information generally comes from secondary uses, many of which are impossible to anticipate at the time of collection, even by data scientists. This is inherent in the dynamism of big data. Novel inferences are its stock in trade.

Propertizers might respond that I could be wildly mistaken about the value of my tangible property as well. I might sell a parcel of remote land for next to nothing, not realizing that the Hilton Corporation was planning to acquire it for its latest world class resort and spa. After all, for any commodity or good, neither potential buyer nor seller can have a perfect idea of its current market value, to say nothing of what it will fetch in five years. But almost all of our decisions are made under some degree of uncertainty. How do I know that my morning coffee won't kill me? Or that my next train ride might not be my last? Still, for most other property or commodities the market is a fairly reliable indicator of value in a way that it systematically is not for personal information. And buyers are likely to be in a far stronger position epistemically than sellers. Whither informed consent?

And another thing: property, intellectual or real, is generally thought to be freely alienable. Says Jessica Litman (2000: 1295–1296), "The raison d'etre of property is alienability: the purpose of property laws is to prescribe the conditions of transfer . . . We deem something property in order to facilitate its transfer." If I sell you my 2005 Civic or the rights to the hit song that I dashed off last month, you can go ahead and offer them to others at whatever price the market commands, and I am implicitly agreeing to these terms at the time of sale. So far, so good. But matters are not so straightforward with personal information. If I sell you information about my last six vacations or the music I have been listening to since September, it looks like I will not be able to stop you from passing the goods along to Jones or Brown, who might turn

around and sell them to Smith or Robinson, who in turn combine the stuff with other bits about me, like the magazines I subscribe to or my commute. Or suppose that I sell you that information so that you can ply me with ads about holiday destinations or music streaming services. Would you be free to use the information for other purposes? If my travel or music tastes suggest a susceptibility to payday loans or for-profit colleges, would the data holder be able to sell that inference (O'Neil, 2016; Samuelson, 2000)? Would there be any way to meter these further uses, or would I simply be out of luck? Moreover, with other types of property, I can form a reasonably good idea of how I will be worse off once I sell it. If you buy my car, I know I'll be riding the bus. If I sell off the chunk of real estate, I won't figure on growing as much corn or wheat next year. And in principle I can buy them back. By contrast, as we saw above, no one, not even a data scientist, is generally able to come to an informed opinion about how surrendering my personal information now will redound to me later. Nor can I plausibly buy the information back. Again, what happened to informed consent?

A further question that propertizers need to answer has to do with enforcement, since no property scheme can exist without it. How would violations be enforced or even be detected? On the one hand, in the vast majority of cases data subjects would not even know that their data was being misused or how this misuse was affecting them, since the harms of big data can be hard to pinpoint. Propertizers might respond with the typical solution for hard to detect crimes like blackmail: up the punishment (Schwartz, 2004). However, whether this measure will be effective beyond the margins is an empirical question, and proponents of propertization need to reckon with the increased fiscal and social costs of more severe penalties or punishments. Also, targets of blackmail know full well that they are the victims! On the other hand, suppose data subjects who have agreed to sell their information obfuscate or deliberately falsify it. Should they be criminally liable? Propertizers need to address these problems.

Finally, I have been speaking glibly of personal information as a form of intellectual property, specifically akin to copyrightable material like books, music, or performances. Intellectual property is non-exclusive. That is, it can be in more than one place at a time. You and I can both have a copy of that catchy new song or the latest page turner. It is also non-rivalrous: One person's use does not affect another's. I can copy your novel or music files and proceed to read or listen to them without depriving you of either. By contrast, if a cupcake is mine, all mine, we obviously cannot both have it, and my eating every last crumb deprives you of the pleasure (Hettinger, 1989; Schwartz, 2004). If I can exclude you from free access to my intellectual creations, it must be for some other reason than that you would be depriving me or others of the ability to enjoy or distribute them. What counts is that without proper protection less intellectual property would be produced and distributed in the first place.

Enter the incentive theory, which I will assume for the sake of discussion is the best justification for copyright in particular and intellectual property generally. The incentive theory denies that creators of intellectual goods like books or music have a moral right to the fruits of their labor. Rather, defenders of the view argue that the law should grant creators or distributors near exclusive legal rights to their products for a period, the copyright term. Doing so provides a motive for creating and distributing those things that others find useful, entertaining, informative, or edifying. So the emphasis, morally speaking, is ultimately on the user as opposed to the producer (Hettinger,1989). The term, ideally, enables creators and distributors to recover their investment and to profit reasonably from their efforts, while excluding others for the time being from helping themselves to the product or churning out copies or knock offs (Hettinger, 1989; Samuelson, 2000). Term length should best conduce to the production of and access to creative products. The ideal term would maximize creation and distribution, consistent with maximum long-term access. The notion is to restrict nearly all free access in the short term to maximize production and access in the long run. Compare declaring a fishery off-limits today to increase yield tomorrow. Any term longer, or shorter, than needed to get these outcomes would be morally questionable.

There is surface plausibility to treating personal information as intellectual property. It too is non-exclusive: Your having access to mine does not rob me of it. It is also non-rivalrous, since your use of it does not deplete my stock. Propertizers might even agree that there should be a limit to how long subjects have near exclusive right to their information, analogous to a copyright term. After all, personal information's usefulness tends to decline over time: A recent record of my Web searches says much more about my present condition and preferences than would a record of them from 20 years ago. Additionally, analogous to fair use, propertizers should be willing to commit some information, like birthdates, to the public domain from the start. The trouble is that personal information is unlike familiar forms of intellectual property—or property in general, for that matter—in

two ways not yet discussed. First, property laws are usually established to protect what is relatively scarce. Yet when it comes to personal information, we have an embarrassment of riches. It is *privacy* that is scarce (Samuelson, 2000). Second, unlike conventional forms of intellectual property, people generally do not need an incentive to create information about themselves. Nor, unlike much intellectual property, do they need to recoup any costs (Samuelson, 2000). Most personal information is generated just by living a 21st-century life. Contrast writing a song or a bit of code, where we can plausibly say that copyright leads to more of the sort being created. It looks like the most eligible justification for intellectual property does not apply to personal information at all. Just what kind of property is it then? I defer to propertizers. Until they can give satisfactory answers to the questions I have raised about their proposal, they have failed to make their case that it can save or substantially protect privacy.

Conclusion

I have tried to show that the main candidates for preserving privacy—notice and choice statements, legislation, obfuscation, and propertization—are inadequate. My next suggestion would be my own solution to the problem, but that I do not have (compare Kripke, 1971). Opting out of the digital grid is not a serious option for most people in the rich world nor for increasingly many in the developing world. If we live a 21st-century life, we will leak data wherever we go, like it or not. Soon the leak becomes a spate. Third parties, good, bad, or indifferent, will be standing by with their analytics to make our lives better or worse. We have seen that privacy can be seriously breached even when people do not volunteer the information themselves, as long as others relevantly like them have. Obfuscation, again, suffers from a similar shortcoming. It would protect a mere slice of our data and do little to secure the narrative drawn from credit cards, toll passes, GPS devices, surveillance cameras, phone apps, and even circulation records or research database activity. Yes, we can routinely swap credit cards, sim cards, create dummy social media accounts, and so on, but will we? And how do we resist the breathtaking colonization of the Internet of things, where dishwashers and espresso machines, for instance, have unique "load signatures," which indicate when they are switched on (Mayer-Schönberger and Cukier, 2013)? It is difficult to imagine any strategy capable of facing down these assaults. It is not that technology is an autonomus, inelucable force whose ravages to privacy are inevitable. Rather, it is that technology and the information flood that it produces will not in fact be stopped, in part because those who benefit from current trends are better financed and organized than the far greater number who stand to lose dearly (Rule, 2007), in part because the vast majority of us either do not care or will remain oblivious to big data's unstinting siege.

This brings me to libraries. Librarians, through their professional organizations, have long championed privacy as a bulwark for intellectual freedom (ALA, 2014; Magi, 2011). Assuring patron privacy—or more specifically confidentiality—promotes free inquiry. In other words, it enhances patrons' autonomy as seekers of information (Rubel, 2014). However, the long-standing confidentiality of circulation records has been partially betrayed by library e-books, particularly those that can be uploaded to a Kindle. As Alan Rubel (2014) points out, this both gives Amazon access to a portion of borrowers' records and also permits the company to merge this information with the substantial chronicle it already has about them. Trina Magi (2013) describes a similar problem with database providers. When users create personal accounts in databases, they are potentially revealing their research interests, with approximately no restrictions on how vendors handle this information. A further problem emerges with journals, where "contracts may have provisions requiring libraries to monitor user activity to detect unauthorized use, and notify publishers" about this (Rubel, 2014: 187). Additionally, as Julie Cohen (1995) pointed out long ago, the move from print to e-journals has given publishers unprecedented access to reader activity and thus to their research interests.

Magi (2013) urges that librarians educate their users about these pitfalls so that they can make "informed choices" and that libraries keep circulation and research profile information in house as far as possible (p. 39; see also Fortier and Burkell, 2015). This is all very well, but will it matter? The same can be said about Rubel's worries about Amazon tracking library users through Kindles. Given the heaps that are already out there, the information gleaned from libraries is trivial. It is like worrying about the breaking of the last barn window when the other 99 are already glass-free.

I see an analogy between the threat to privacy and the challenge of climate change. My guess regarding the latter is that we will not get the kind of international cooperation needed to stop the worst havoc, given inertia and the powerful forces that have an interest in sticking with fossil fuel. Likewise with privacy: Over time, our privacy "immune systems,"

to use James Rule's (2007: 165) metaphor, grow weaker; threats to privacy tend to encounter less resistance as the years go by. Imagine telling someone 40 years ago that today every other person on the planet would choose to tote a glorified tracker around with them at all times! Does anyone really think, for instance, that privacy concerns will be foremost in the design and deployment of self-driving cars?

At the turn of the century *The Economist* (1999: 15) shrewdly speculated that:

> All . . . efforts to hold back the rising tide of electronic intrusion into privacy will fail. . . . Twenty years hence most people will find that the privacy they take for granted today will be just as elusive as the privacy of the 1970s seems today. . . . Many might prefer to eschew even the huge benefits that the new information economy promises. But they will not, in practice, be offered the choice.

Evidently. Like it or not, the time has come to give up the ghost of privacy and thus call off the moral debate to save or restore it. Anyway, people evidently enjoy the convenience that big data and its analytics have offered them in terms of movie or vacation recommendations, location services, easy contact with friends and acquaintances, and so on. To some extent, people have acquiesced in the demise of their own privacy. Defenders of privacy need to deal with the fact that it just might be that people *like* being profiled. Also, I mentioned at the outset that big data is transforming the social sciences and our approach to public health. Maybe people will someday view the end of privacy the way we think today about the loss of innocence after Copernicus or Darwin. They could well decide that trading away privacy was worth it in the light of the very considerable benefits that big data has offered for our understanding of ourselves, the control of disease, the efficiency of smart public transportation, and the safety of autonomous cars. Or they might never know what they are missing. Perhaps our real concern should not be with privacy but with the widening gap in wealth and power that big data seems to be driving. *That* is anything but inevitable.[3]

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. I would like to thank an anonymous reviewer for prompting me to make this clarification.
2. I would like to thank Phil Swan for this point.
3. I would like to thank John Buschman, Jane Carter, Don Fallis, Shannon Oltmann, and Philip Swan for their feedback. I would also like to thank attendees of the 2017 Information Ethics Roundtable, held at the School of Information Sciences at the Univerity of Illinios, Urbana-Champaign; and attendees of the 2017 annual meeting of the International Association of Computing and Philosophy, held at Stanford University.

References

Acquisti A (2014) The economics and behavioral economics of privacy. In: Lane J, et al. (eds) *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. New York: Cambridge University Press, pp. 76–95.

Acquisti A, Taylor C and Wagman L (2016) The economics of privacy. *Journal of Economic Literature* 54(2): 442–492.

ALA (American Library Association) (2014) Privacy: An interpretation of the Library Bill of Rights. Available at http://www.ala.org/advocacy/intfreedom/librarybill/interpretations/privacy (accessed 13 December 2017).

Angwin J (2014) *Dragnet Nation: A Quest for Privacy, Security, and Freedom in a World of Relentless Surveillance*. New York: Times Books, Henry Holt and Company.

Angwin J, Scheiber N and Tobin A (2017) Facebook job ads raise concerns about age discrimination. *The New York Times*, 20 December, p. 1.

Barocas S and Nissenbaum H (2014a) Big data's end run around anonymity and consent. In: Lane J, et al. (eds) *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. New York: Cambridge University Press, pp. 44–75.

Barocas S and Nissenbaum H (2014b) Big data's end run around procedural privacy protections. *Communications of the ACM* 57(11): 31–33.

Bilton N (2017) *American Kingpin: The Epic Hunt for the Criminal Mastermind behind the Silk Road*. New York: Penguin.

Brunton F and Nissenbaum H (2011) Vernacular resistance to data collection and analysis: A political theory of obfuscation. *First Monday* 16(5).

Brunton F and Nissenbaum H (2013) Political and ethical perspectives on data obfuscation. In: Hildebrandt H and De Vries K (eds) *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology*. New York: Routledge, pp. 164–188.

Brunton F and Nissenbaum H (2015) *Obfuscation: A User's Guide for Privacy and Protest*. Cambridge, MA: MIT Press.

Cohen J (1995) A right to read anonymously: A closer look at copyright management in cyberspace. *Connecticut Law Review* 28: 981–1040.

Cohen J (2000) Examined lives: Information privacy and the subject as object. *Stanford Law Review* 52(5): 1373–1438.

Duhigg C (2009) What does your credit-card company know about you? *New York Times Magazine*, 17 May, pp. 40–45.

Economist (1999) The end of privacy. *Economist*, 1 May, pp. 15–16.

Fortier A and Burkell J (2015) Hidden online surveillance: What librarians should know to protect their own privacy and that of their patrons. *Information Technology & Libraries* 34(3): 59–72.

Gandy O (1993) *The Panoptic Sort: A Political Economy of Personal Information*. Boulder, CO: Westview Press.

Hettinger E (1989) Justifying intellectual property. *Philosophy & Public Affairs* 18(1): 31–52.

Howe D and Nissenbaum H (2009) Trackmenot: Resisting surveillance in web searches. In: Kerr I, Lucock C and Steeves V (eds) *Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Networked Society*. Oxford: Oxford University Press, pp. 418–436.

Kripke S (1977) Identity and necessity. In: Munitz M (ed.) *Identity and Individuation*. New York: New York University Press, pp. 135–164.

Landau S (2015) Control use of data to protect privacy. *Science* 347(6221): 504–506.

Laudon K (1996) Markets and privacy. *Communications of the ACM* 39(9): 92–104.

Lessig L (2002) Privacy as property. *Social Research* 69(1): 247–269.

Litman J (2000) Information privacy/information property. *Stanford Law Review* 52(5): 1283–1313.

Magi T (2011) Fourteen reasons privacy matters: A multidisciplinary review of scholarly literature. *Library Quarterly* 81(2): 187–209.

Magi T (2013) A fresh look at privacy: Why does it matter, who cares, and what should librarians do about it. *Indiana Libraries* 32(1): 37–41.

Mai J (2016) Big data privacy: The datafication of personal information. *The Information Society* 32(3): 192–199.

Mayer-Schönberger V and Cukier K (2013) *Big Data: A Revolution that Will Transform How We Live, Work, and Think*. Boston, MA and New York: Dolan/Houghton Mifflin Harcourt.

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

Nissenbaum H (2011) A contextual approach to privacy online. *Daedalus* 140(4): 32–48.

Oberhauser K (2011) Monarch butterfly. In: *Environmental Encyclopedia. Vol. 2*. 4th edn. Detroit, IL: Gale, pp. 1091–1093.

Ohm P (2014) Changing the rules: General principles for data use and analysis. In: Lane J, et al. (eds) *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. New York: Cambridge University Press, pp. 96–111.

O'Neil C (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Pekala S (2017) Privacy and user experience in 21st century library discovery. *Information Technology & Libraries* 36(2): 48–58.

Rubel A (2014) Libraries, electronic resources, and privacy: The case for positive intellectual freedom. *Library Quarterly* 84(2): 183–208.

Rule J (2004) Toward strong privacy: Values, markets, mechanisms, and institutions. *University of Toronto Law Journal* 54(2): 183–225.

Rule J (2007) *Privacy in Peril*. New York: Oxford University Press.

Samuelson P (2000) Privacy as intellectual property? *Stanford Law Review* 52(5): 1125–1173.

Schneier B (2015) *Data and Goliath: The Hidden Battles to Capture Your Data and Control Your World*. New York: WW Norton.

Schnur D (2002) Mimicry. In: Cobb A (ed.) *Animal Sciences*. Vol. 3. New York: Macmillan Reference USA, pp. 121–123.

Schwartz P (2004) Property, privacy, and personal data. *Harvard Law Review* 117(7): 2055–2128.

Singer N (2013) A data broker offers a peek behind the curtain. *The New York Times*, 1 September, p. 1.

Steel E and Angwin J (2010) On the web's cutting edge, anonymity in name only. *Wall Street Journal*, 4 August, A1.

Stephens-Davidowitz S (2017) *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us about Who We Really Are*. New York: Dey St.

Tavani H and Moor J (2001) Privacy protection, control of information, and privacy-enhancing technologies. *Computers and Society* 31(1): 6–11.

Walker J (2013) Data mining to recruit sick people. *The Wall Street Journal*, 17 December, B1.

Author biography

Tony Doyle is a Reference/Instruction Librarian and Associate Professor at Hunter College (CUNY) in New York City. He is also an Adjunct Associate Professor in the Philosophy Department at Hunter. He has an MLS from Queens College (CUNY) and a Master's in Philosophy from Northern Illinois University. His research interest is in the ethics of privacy, focusing on how privacy has been affected by digital technology.