Inferences are many: protecting privacy beyond definition and disclosure

Matt Voigts[1]

[1] PhD candidate at Horizon Centre for Doctoral Training, University of Nottingham, Jubilee Campus, Nottingham, UK, NG8 1BB, matt.voigts@nottingham.ac.uk
PhD studies partnered with the Open Rights Group, London, UK

1 Introduction

Theoretical literature often discusses informational privacy as being managed via the disclosure of discrete, definite items of personal data. In this view, information is either known or not known to various parties, and individuals maintain agency over information by limiting its transmission. This paper seeks to challenge these assumptions by conceiving privacy as strongly reliant on inferential processes of meaning-making.

In practice, data is rarely unambiguous, but rather an instantiation of a particular meaning. In the case of the information privacy is invoked to protect, our personal identities, beliefs and health records are often complex and fluid. Likewise, data itself is often not a simply 'true or 'known'. Privacy conventions furthermore often prioritize maintaining order in the public sphere over individual agency. In the global digital infrastructure, data is moved across contexts as a technical matter of course, making disclosure and context difficult to definitively know. The data analytics process, moreover, is continually being refined, and often abstract by human standards. This makes it difficult to relate disclosure to subsequent meanings produced, and in some cases may undermine individual personal agency that privacy has been valued to safeguard.

After exploring these challenges, this paper seeks to address them by situating information and its disclosure in a wider process of meaning-making activity, and the many valuable insights of privacy literature within a wider body of socio-cultural, technical, and anthropological literature involving social meaning-making. This view suggests that much of the over-reaching emphases on definition and disclosure result from an idealized vision of personal control that was neither the historic nor contemporary global norm. To be efficacious in protecting the personal agency that privacy safeguards – beyond regulating what information is *allowed* in the public sphere and private databases – we must foreground a careful consideration of the status we accord the meanings produced with information.

2 Disclosure and Definition in Informational Privacy

This section will discuss informational privacy as conceptualized by name across a variety of disciplines, including the ways in which it is invoked to protect personal agency. In the process, it will illuminate two underlying assumptions that suffuse some of the literature: that information is definite, and its flow can be controlled by limiting disclosure of information.

Privacy has been considered by scholars across disciplines as agency in controlling access to oneself, including information about oneself. Warren and Brandies' 1890 articulation of the concept as "the right to be left alone" remains an oft-cited reference point for the concept [1]. In more recent theory, the philosopher Marmor defined privacy as "a reasonable measure of control over the ways in which they can present themselves (and what is theirs) to others" [2, p.4]. Westin – whose 1967 *Privacy and Freedom* [3] remains a landmark text, on the topic, and who for several decades produced indexes on American attitudes toward the topic – summarized privacy as "the claim of an individual to determine what information about himself or herself should be known to others." [4, p. 431]. Drawing strongly from the American legal tradition, Nissenbaum described privacy as "contextual integrity", proscribing that information should be disclosed according to normative expectations; information intended for particular contexts should not necessarily flow into others [5][6][7]. Drawing from an extensive literature review, Koops et al created a taxonomy of eight ideal types of privacy, over which individuals may exercise influence: bodily, spatial, communicational, proprietary, intellectual, decisional, associational, and behavioral [8, p. 69-70). They further overlay a ninth concept of informational privacy, suggesting control of information is the most dominant factor in managing access to the self.

Many of the arguments in favor of privacy argue for its value in preserving individual agency. Frequent reasons offered to value privacy include: preserving the ability to present different aspects of oneself in different situations [5] [6] [7]; encouraging democratic participation by allowing individuals to express controversial ideas and organize politically without fear of retribution [9][10][11][12]; protecting one's internal sense of identity [8, p.38]; and negotiating and delineating one's intimacy with other. Anderson and Dourish suggest that sharing and withholding information "cement[s] a bond between those who share it and mark their difference from those with whom it is not shared," just as learning what is and isn't off-limits to ask within a social group may encourage group cohesion [13, p. 332-3]. The information privacy protects is personal, while the consequences of its propagation may be both personal and social. Given privacy's intimate connection to the social, however, individual agency is not unlimited. As Post writes, "most persons desire to define themselves and to have others accept their self-definition. But this desire is incompatible with the ways in which public discussion necessarily appropriates the authority and the power to define persons that are the subject of public consideration." [14, p. 2091].

The change that digital data – and 'Big Data' analytics – have wrought often described in terms of increasing information's scale, of expanding its 'volume, variety, velocity, and variety' [15] [16]. The enhanced reach and longevity of information has

profound implications for privacy. Marwick and boyd describe the anxiety of different social expectations converging in a single platform as "context collapse" [17]. Solove [18] and Ronson [19] offer a wealth of accounts of individuals whose publicly-controversial actions live in indefinite Internet infamy due to the longevity of online data. The negative consequences are often described in terms of personal stress, social stigma, and professional sanction (IE, getting fired).

Saussure's distinction between the *signifier* and the *signified* recognized that meaning-making is an action of the interpreter; we cannot – in a technical sense – 'make' others hear what we intend [20]. However, privacy theory suggests that individuals act within cultural contexts in which norms and behaviors that allow them to predict how their actions will be received. This is important to preserving privacy because, as Marmor argues, "I can only make choices about what I reveal to others if I can predict the casual relations between my conduct and others' uptake" [2, p.12]. Privacy, according to DeCew, "is not merely limited to control over information. Our ability to control both information and access to us allows us to control our relationship with others." [21] That right to calibrate one's conduct to expectations in a variety of situations with conflicting norms is an important justification for contextual integrity, with Nissenbaum arguing for "transparency and choice" as being an important control mechanism for gauging what to disclose to whom [7]. International ethnographic research confirms that people exercise control within their understandings on more 'public' social media platforms (such as Facebook, where people are "friends" with people from many parts of their lives), individuals generally behave conservatively, calibrating their activities to the contextual norms of the offline lives [22]. In more 'private' social media (such as WhatsApp), more private norms may apply [22] [23].

For many discussions, Warren and Brandies' "right to be left alone" (or, indeed, any of the above) may suffice as precise enough to capture a sense of privacy: some combination of agency over access to oneself and the right to disengage with others at one's choosing. Yet, as with many topics, the ultimate amount of ink devoted suggests difficulty in capturing its nuances effectively. As Post wrote, "privacy is a value so complex, so entangled in competing and contradictory dimensions, so engorged with various and distinct meanings, that I sometimes despair whether it can be usefully addressed at all" [14, p.2087]. At the same time, the quantity of the literature also makes it difficult to suggest that *privacy* has constituted a cohesive discussion. As such, I offer my claims with the caveat that they represent my interpretation of general trends rather than exist as exhaustive absolutes.

Nonetheless, below I will expand the discussion to argue that they largely portray information as *definite* in the sense that it has unambiguous, readily-understood meanings. It becomes problematic when it leaves particular contexts, not because it has become the building blocks for new meanings. The primary mechanism of context change is *disclosure* – that is, regulating privacy involves controlling under what circumstances information is revealed to whom

Examples in the above literature frequently focus on binaries of revelation. Someone may or may not enter another's home. Law enforcement may or may not stockpile

mobile phone meta-data providing a record of who one called, when, and from where they placed the calls. Marmor is exceptionally rich in describing commonplace scenarios of observation and transmission, including airport body scans, CCTV surveillance, and a teacher pseudonymously posting an anecdote about a student. "Suppose that I just walk through the streets of downtown Los Angeles…" he hypothesizes, "and later find a photo of myself taken on the street posted on the Internet, available for millions to see…" [2, p. 20]. He does not rule that all these scenarios are privacy violations. If a violation does occur, however, it is usually via some combination of observation and transmission outside social norms. Airport security might or might see me in a degree of undress and vulnerability through a scanner of defined observational abilities; a teacher may or may not post an anecdote, from which the student may or may not be identified.

This is consistent with Nissenbaum's contextual integrity: violations occur when transmission has gone awry, or overly intrusive observations are made contrary to reasonable expectations. Nissenbaum's original 2004 piece on 'contextual integrity' was written amid concerns of surveillance following 9/11, well into the digital era [5]. In a 2011 piece, she directly applied her theory to online data, concerns generally involve transmission – as when companies sell digital data to other companies [7]. Both however, share Marmor's informational epistemology in primarily considering these pieces of data as having fixed meanings, and disclosure as being a primary challenge to individuals' rights. In academic and popular literature, concerns about data analytics prioritize controlling data collection as a prerequisite to avoid subsequent consequences [9] [10] [11] [12]. Koops et al state that privacy may govern "subsequent control after access has been granted." [8, p.67], on which they do not extensively elaborate. They do, however, mention a Czech legal provision involving (their words) "letters that have been delivered and are subsequently stored" [p.37]. This suggests a concern based, again, around a model of a series of discrete disclosures.

3 – Problems definition poses for individual agency

Above, I described why privacy is valued (broadly: that out-of-context information and undue exposure can have unfortunate consequences) and how theory suggests we protect it (broadly: via individuals exercising informed agency over their personal information). In this section, I will argue that not accounting for the interpretative aspects of information can undermine individuals' control over it. Suggesting we follow contextual norms of transmission involves defining what those norms are; consciously portraying a vision of one's life one wishes to project requires making definitive decisions about (and drawing attention to) personal information that is often indefinite and ambiguous. At the same time, social conventions about what it is and isn't appropriate often prioritize social cohesion over individual agency. In seeking to grant individuals control of information, we may inadvertently encourage them to inscribe heretofore indefinite concepts in social structures which are biased against individual agency.

An example of the complicated relationships between meaning-making and disclosure – and how they can work at odds with one another – can be found in the "Streisand

Effect." The effect is named for an unsuccessful legal action in which entertainer Barbara Streisand attempted to suppress a photo of her seaside mansion as part of a larger photography project seeking to document coastal erosion. Streisand claimed the coastal photos invaded her privacy, and her subsequent legal action brought more attention to the photos than they likely would have received otherwise - including having an Internet phenomena named from the incident. The photo project's website extensively documents the coverage the lawsuit and 'Effect' have received [24].

In the Streisand case, seeking to obscure information inadvertently publicized it. Moreover, however, it demonstrates that information and context weren't the sole matters disputed matters, but that they were bound up in the public *meaning* of the information. Before Streisand sued, the photo was positioned as a document of coastal erosion, out of which one could incidentally view an estate. In one interpretation, we could take at face value the claim of Streisand's legal action, that it violated her privacy. Any individual has an interest in remaining obscure at times, and a celebrity of her status may have a particular interest in protecting her privacy from obsessive fans or other ill-wishers. The role of the celebrity in American culture involves many complex decisions about self-presentation in public and media. Yet for this reason, one could also read the action as partially being a strategic decision about wishing the extent of her apparent wealth be deemphasized from her public image – or perhaps a petty attempt at exercising power for its own sake. Given her reputation as a supporter of environmentalist causes, perhaps it was a stealth plan to draw attention to coastal erosion. Regardless of how one interprets it, the case didn't just call attention to the photograph and make it a contested battleground of meaning, it forced the photo to take mean *something*. Disclsoure implies definition.

The encouragement data provides to publicly define otherwise undefined or ambiguous topics may be a source of anxiety unto itself. Miller suggests that tensions over the anxieties over privacy might be particularly pronounced in England, stating that a "good deal that the English see as characteristic of being English has to do with the complex relationship between public and private." [25, p.5]. He argues that English use of Facebook involves a "Goldilocks strategy" of "middling distancing": keeping social relations at just the right distance, just as in the fairy tale, Goldilocks sought an appropriately-sized bed [p.100]. In this understanding, anxieties over privacy on social media is also an anxiety over *definitions themselves* and how social media forces users to sort their social relationships into categories like 'friend lists'. He writes that "social media simply makes evident the exquisite sensibility of many English people to the exact state of their relationships" [p.109]. Miller et al discuss how anxieties over defining relationships on social media play out in nine cultural contexts [22].

More fundamentally, this defined information – and the structure in which it exists – is never neutral. It is always created to specific parameters, and these parameters carry the imprints of their emphases, prejudices, and interests. As Bowker and Starr describe, data sets are records are made with purposes that can only capture (at best) a fraction of events as they happened [26]. As Graeber discusses, documentation that is official and bureaucratic-minded (such as birth records, passports, legal documentation and financial records) have an official status and air of veracity beyond records

created by individuals with less institutional power, and reflect the concerns of the more powerful interests [27]. Until relatively recently historically, such records were often the most prominently (or only) created and/or stored records of individuals' lives and actions. The digital public sphere, however, has democratized activities of meaning-making and information documentation, at least among those with the technical resources and skills to make Facebook posts and Tweets.

With this in mind, there at least three exceptional challenges the act of definition poses to personal agency. One, if one wishes to respect the norms of privacy (as in contextual integrity) along the concerns of individuals, one must define them. In doing so, one must also reckon with the fact that individuals and institutions often misjudge risks around the topic. More people describe concerns about privacy than take action to digitally safeguard their own [28]. The amount of national defense resources that have been invested in security to protect against terrorism stand in contrast to the more statistically-probable, addressable safety issues in Euro-American countries such as preventable diseases, nutrition, and car accidents.

Two, despite how often it is categorized, few types of information may be as poorly-suited to the categories of ready-made spread-sheets as personal information. One's understanding of (for example) one's own sexual orientation may be difficult to easily categorize and change over one's lifetime. Health – a frequent domain of privacy contentions – is often encoded by medical professionals in orderly-looking spread-sheets. But in practice such claims often prioritize the concerns of the system over strict fidelity to 'what happened'. Bowker and Star's work on categorization was in large part based on health records. Medical terms themselves are socio-cultural vectors of meaning. In practice, as Mol's ethnography explores, diagnoses do not necessarily capture doctors' or patients' lived experiences, and even 'definitive' diagnoses may remain open-ended [29]. Their coding reflects a categorization of events that happened rather than a strict record of 'reality', as demonstrated in the complexities, inaccuracies, and inconsistencies in the death records Bowker and Star discuss and Rampatige et al statistically explore [30]. As Onouhu phrases the issue, "as we abstract the world, we prioritize abstractions of the world" – that is, as we encode data to particular parameters, we come to view the world through these abstracted parameters. With digital data, the various abstractions layer and build on one another [31].

Three, social conventions around privacy often prioritize social stability over the individual's influence in the public sphere. As Marmor notes, in social conventions regarding propriety and obscenity, it is "the *public* zone that is in need of some protection, not the private. People have legitimate expectations about what they encounter in public spaces." [2, p.24]. Kulick describes how classifying sexuality as 'private' in Sweden precluded disabled individuals under institutional care from accessing related services and information [32]. Likewise, the norms of public / private division do not take into account the intricacies of family dynamics. Domestic violence was long held to be a 'private matter', sheltering its perpetrators and perpetuating its practice. One can see the subtext of prioritizing one's value as an instrument of labor in some concerns about 'privacy' online: would one want a potential employer to see a post?

4 – Disclosure is a situated concept

To summarize the above, aspects of privacy theory emphasize managing the disclosure of discrete items of definite information. In the process, this may force otherwise ambiguous concepts into defined data points, in the process complicating individual agency. This section suggests affording more attention to how information is publically instantiated, interpreted and accorded social validation – all important to the agency over personal information that privacy is invoked to protect, and elements that set the circumstances for inferential thinking in the public sphere. This involves reclaiming privacy from idealized visions of personal control, as well as situating disclosure within broader processes by which information is afforded attention.

The first act of situating the concept, then, is to question our concerns about privacy. It is important that privacy theory does not proceed from an idealized image of control. Squire describes all secrets as "partial" in the sense that naming them implies the possibility of their discovery [33]. The notion that private information can be definitively controlled may be an especially strong notion in the globally atypical, Euro-American societies in which much privacy scholarship is produced. These are sometimes referred to by the shorthand "WEIRD": Western, Educated, Industrialized, Rich, and Democratic [34**].** Simmel grounded "the increasing significance of privacy in the fragmentary character of modern [Western] individuality, and account for this fragmentation in terms of transformation of the pattern of group-affiliation." [36, p. 330] In this view, the desire for contextual integrity arose with modernity, as people began affiliating themselves in multiple contexts that involved potentially conflicting roles. Historically and presently, most people's control over the space around them in many situations (included the home) has been limited. In the history of Europe alone, there are many examples of common present and historic living arrangements in which 'privacy' controls are limited to the individual: monasteries, prisons, tenement buildings, hostels, university dormitories, nursing homes, hospitals, and the semi-detached house in which I live, whose thin walls provide a scant auditory border with my neighbors.

As such, I would caution against phrasing privacy in terms of how Human-Computer Interaction practioners Dourish and Anderson evaluated privacy as "a catch-all term for how individuals might lose control of information" [13, p. 322], which differs from above-quoted definitions of privacy in that it directly suggests that individuals by default *possess* information. These claims to the normativity of control also underpin parts of Nissenbaum's arguments, as when she suggests there exists a:

> "simple and ages-old idea of the sanctity of certain spaces…For example, "a man's home is his castle"—a person is sovereign in her own domain. Except when there are strong countervailing claims to the contrary, this principle apparently endorses a presumption in favor of people shielding themselves from the gaze of others when they are inside their own private places." [4, p.129-30].

It would be more accurate to say that – despite how the claim to space may appear "ages-old" – individual rights to such spaces have been heavily filtered through the post- 18th century conceptions of property ownership she subsequently cites, including Warren and Brandies. In that original legal paper, the authors argued that privacy was as an existing aspect of American proprietary rights, but one that needed to be distinctly clarified with respect to information as a result of the rise of what was then 'new media': photography and newspaper gossip columns [18, p.105-110].

These pre-digital proprietary precedents – such as U.S. Constitutional rights against undue search and seizure and soldier-quartering – materially differ from the current information landscape and its transmission possibilities. A horse or house can be singularly owned. Information always maintains a material base, be it a scrap of paper, a hard drive [36] the misleadingly-named 'cloud' [37], or the human brain. Digital data can, however, be replicated much easier than a horse or plot of land. In fact, digital data is replicated repeatedly in the course of its journey from a server to an individual's computer. The path data makes varies, but can be determined with a traceroute. 'Privacy by design' projects seek to build infrastructures that protect data by regulating its technical workings and affordances, but at present, the global infrastructure requires data be 'disclosed' and inscribed many times on its route through the global infrastructure.

The digital infrastructure, however, is designed in such a way as to distract from its physicality. Digital data feels both immaterial and definite, and yet – in many respects – is neither. As Kirschenbaum argues, "computers are unique in the history of writing technologies in that they present a premediated material environment built and engineered to propagate an illusion of immaterialty" [38, p. 135]. This has practical consequences for individuals' conceptions of privacy. Schneier suggests that "we tolerate a level of electronic surveillance online that we would never allow in the physical world, because it's not obvious or advertised" [12, p.33]. It also (as will be discussed later) poses challenges for contextual integrity – if we're to govern action by the norms of transmission, what do we do when the human expectations for transmission are radically different from the technical means by which they are achieved?

The idealization of contextual control should not be made the enemy of self-expression regarding personal information. As mentioned, Miller et al discuss an emergent global 'conservatism' on public-facing social media [22]. Yet if we would say this is felt as undue pressure, one should also note that *all* situations have social expectations. It is difficult to imagine behavior that is *entirely* divorced of expected consequence or positioning. The emergent conservatism of Facebook is perhaps more akin to the conservatism one experiences in physical public spaces, such as a restaurant. Second, this conservatism exists in light of 'polymedia' [39] [40] in which individuals use a variety of mediums on a spectrum of public to private. People may choose to express certain ideas on a more private social medium, such as messenger services like WhatsApp, thereby taking agency over their decisions within available options, whatever imperfect controls they have over the infrastructures that facilitate their communication.

To summarize, individuated personal information control is highly situated, dependent on a number of variables, and a relative historical rarity. The digital infrastructure relies of copying and transmission of information. In such an environment, it makes sense to situate disclosure within a larger set of conventions and controls that govern how we afford information attention.

Selinger and Hartzog (from philosophy and law, respectively), argue that strict categorical 'privacy' may be less an expectation for information than "obscurity", "the idea that information is safe – at least to some degree – when it is hard to obtain or understand" [41 p. 2]. The philosopher Nagel discusses a wide range of "concealment and exposure" behaviors as part of a variety of behaviors governing the public and private. He notes that "[c]oncealment includes not only secrecy and deception, but also reticence and nonacknowledgment. There is much more going on inside us all the time than we are willing to express, and civilization would be impossible if we could all read each other's minds." [42]. At the extreme individual, cognitive level, the anthropologist Jackson describes how people cycle between periods of focused attention and withdrawal, in effect seeking solitude moment by moment [43, p.1-21]. At the observer's end, Goffman described a pre-condition for privacy as "civil inattention" [44], which Koops et al elaborate as the "norms of seeing but not taking notice (or perhaps rather, demonstrating not to take notice), for instance by averting one's eyes." [8, p.58].

There is a case to be made to go back to a limited definition of privacy that describes the disclosure of data alone, with other labels applied to other parts of the information control chain. Regardless, according greater emphases to these elements – concealment, obscurity, and other ways in which people choose to ignore personal information, and the choices they make within available possibilities – suggests a more complex process by which attention is granted, and moves away from a proprietary conception of privacy as involving owned, defined information.

5 – Inferences and meaning-making

Having discussed how neither personal information nor disclosure is definite, this section describes the underlying challenge for personal informational agency as involving how all these indefinite elements are valued in the public sphere. Claims *based on* information are made through inferential processes, rather than discrete items of data itself. In digital analytics, the meanings are frequently created via processes that humans would consider abstract. In order to protect personal agency in information, privacy must address not just the meanings, but the roles in the process to arrive at them.

Unlike a mythic expectation of information control, the role of inference in the public sphere is not only entwined with the digital age and Western modernity, and notably, it occurs even in situations in which people live in close spatial proximity over which their control is not individuated. An anthropological examples of discussions pertaining to privacy comes in Stang's *A Walk to the River in Amazonia* [45]. There is no digital technology use (let alone electricity) described for the Mehinaku tribe the book

describes, and little privacy as it has been heretofore discussed. Stang writes: "There is an intense closeness of social life for the Mehinaku. In the houses, extended families of ten to fifteen or so people live in a space without walls, moving quite freely between the hammocks, so there is little spatial privacy at all…To be alone…is dangerous." [p.165]. This is not to suggest that there are no norms of space usage; certain areas are gender-segregated, for example, and men and women use certain spaces at different times of the day to conform to conventions. Yet even with people who live in close proximity, there is still a substantial amount of interpretive room built into social relations. Gossip is common, as are supernatural explanations for observable events. She writes: "I was shocked by the things some people would whisper to me about others they seemed very close to…At one point I realized that a man from every single house had been accused of sorcery to me." [p.167].

While Stang does not speculate on the connection between space and interpretation, I suggest that the Mehinaku's practices suggest that even in areas with little spatial privacy (or expectation thereof), speculative and interpretative spaces in regards to information remain open. This is further supported by the Mehinaku's cosmology which forms a substantial topic of discussion in her book. The group believes most objects and living things they encounter in nature is a *copy* of a true form, and that different material realities are perceived by those in different relations to the spirit world. For the Mehinaku, everyday life involves perceiving elements of the world as material copies of an ideal form, and understanding that most things one encounters has a potential to be experienced very differently from person to person. By contrast, digital technology users live in a world where copies and interpretation are manifestly part of a vast material technological infrastructure, but one that's myth discourages us from perceiving its material nature and processes of disclosure, copying, transmission and inscription.

Stang's account of the Mehinaku suggests that even within close relationships, people develop interpretive room. At the same time, people in close relationships share large quantities of information with one another, deliberately and incidentally, and develop complex, nuanced understandings of context. Among Nissenbaum's examples, she suggests that a "sexual partner may be entitled to information about the other's HIV status, although the same demand by a friend is probably not warranted." [5, p.143] Yet while it may be unfair to *ask* such a thing of a friend, Squire found that HIV-positive individuals who *hadn't* disclosed their status inferred whether friends and family knew based on the precedents of their friends' and families' behavior. This process is less frequently rooted in a defined 'smoking gun' disclosure (such as a parent outright stating 'I know') than the intuitive, intimate understandings that close social relationships foster – based in information, but not necessarily directly on a definitive disclosure [33]. It is likely that those who come to believe their friends may be HIV-positive employ a similar process to arrive at their understanding.

In the above, whether a friend sees a scrap of paper from a clinic or 'puts together the pieces' from a variety of incidental material, the end result of understanding may be similar: someone comes to *know* (or at least believe) that their friend has HIV. This is where privacy's intense focus on disclosure betrays an imbalance in what inferentially

happens to the information *after* it has been disclosed: if it is of tremendous importance and interest for individuals to control where their (instantiations) of personal information go, what are they to make of how others may *work out* meanings on their own? Is there a contextual barrier on what is acceptable to devote one's analytical capabilities? How much thought can be put in to determining what one friend ate for dinner the previous night before such intrusions are contextually violating?

Does it matter if one's friend is Sherlock Holmes? Is the burden upon Dr. Watson to simply accept that no matter how focused he thinks his words are, how little he says, how thoroughly he cleans the 221B Baker Street flat they share, the world's most famous detective will simply be able to work out every intimate detail of his life, and this is simply a factor for which Holmes' acquaintances must account? Is the burden on Sherlock Holmes to occasionally switch off his reasoning power (if he is capable of such a thing)? Does it matter if Holmes considers his reasoning process as 'deductive' – as simply following a logical chain of events – while readers of his exploits may consider it 'inductive' – as applying small observations to a larger theory?

In digital analytics, however, disclosure of data is noticeably asymmetrical from the meanings that can be drawn from them in important ways that make maintaining a sense of uptake difficult, likely more so than in the interpersonal interactions of everyday life. On one level, data analytics is designed to generate meanings, and technology is increasing their ability to generate meanings. As Onouhu phrases it, "especially combined, data sets reveal far more than intended" [31]. The meanings that can be drawn from a data set – especially as they increase in size, and are combined with other data sets – increase in ways that are difficult to predict. What analytics may generate tomorrow may be different from what they produce today.

Two, the process of analytics is different from the process of human uptake. Researchers and data analysts are removed from the contextual references points from which friends might be able to draw inferences. In the contextual terms interpersonal communication, the gulfs between *signifier* and *signified* may be particularly pronounced in data analytics. The meanings analytic processes generate are often indirectly or hesitantly associated with what info users are expecting to convey publicly, and – in the case of meta-data – are often the byproducts of more direct communicative activity. If we use Squire's term of the 'partial secret', it is not just that analytics may uncover things we didn't *think* we were revealing, it's that it can instantiate meanings we didn't even consider. Analytics may be unconcerned with the image of your lunch salad you posted. It may be tremendously interested in login times, geolocation data, your mobile phone's model, whether or not you have the most recent software updates – and how it can use these things to either generate meanings in and of themselves, largely unrelated to salad, about your HIV status. Or it could be applied to use these items of information to identify you enough to link that information to other data sets – and thereby increase the meaning-generating capacity.

For a particularly noteworthy example of what analytics can suggest about personal information, Kosinksi, Stillwell and Graepel analyzed 58,000 volunteers' Facebook likes to predict "highly sensitive personal attributes including: sexual orientation,

ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender" [46]. As an example, these claims are different from the discrete and disclosable 'information' that privacy often assumes in at least three ways. In regards to sexual orientation:

1. The conclusions do not make a definite claim that any individual is (for example) homosexual, merely that the analytic methods have been able to determine homosexuality with an 88% accuracy rate.
2. The conclusions were likely not based on individuals declaring their sexual orientation direction (as in, stating "my sexual orientation is ____") , but inferred based on correlations pertaining to 'liking' activities that may or may not have been directly or indirectly related to sexuality by human-discernable understandings.
3. The conclusions may or may not align with how individuals see themselves. The claims may outright conflict with their self-understanding, or be inapplicable to those who see themselves outside of the study's imposed definitional categories.

Point one suggests that – unlike a tidy, discrete disclosure – the information produced is an instantiation of an indefinite claim, here made for the purposes of proof-of-concept research. To apply contextual integrity to the second point, the extent to which revelation may be connected to uptake is a question of the degree to which one regards digital inferential process as similar to the human inferential processes. Could a human have expected that their liking activity could be reasonably used to determine their sexuality? The third point, however, is a particular challenge for the burden it places on inter-personal narrative. Does this suggest that the individual has primacy of interpretation of their data, or that the analytic process could *understand* the individual on a deeper level than s/he understands him/herself? How is this inference to be regarded by different publics? What do we make of the fact that someone has instantiated it as a meaning in the first place?

The answers to how to regard this data publicly are unlikely to come easy. However, in identifying them – and according more attention to the interpretive processes that arrived at them – we cede less individual agency to the terms of the analyst. Despite the neat and tidy appearance of data, outputs are arrived at through processes with interests, biases, and oversights. Limiting disclosure only provides a measure of protection, and according inferentially-produced data an inflated truth value ultimately takes away from individual's capacity to position their actions for interpretation within their immediate contexts.

These are ultimately matters of social values, of which sheltering the disclosure of actions offers limited protections after a certain point. Schneier suggests that privacy may be an important incubator for social liberalization by allowing stigmatized practices to gain popular support before they receive official sanction [12] – as, for example, homosexuality was widely extant in the UK before homosexual practices were

decriminalized in 1967. What would the analytic ability to work out sexual orientation, with an 88% success rate, mattered if it were developed and legally implemented there in the early 1960s? What would it effect today, if Kosinksi et al's methods are applied in countries today where homosexual acts may be punished by the death penalty? Privacy may protect actions that are acceptable in certain contexts and stigmatized in others, but it does not resolve the oft-consequent questions of what is considered acceptable, valued and accorded attention interpersonally.

5 Conclusion

Current theoretical conceptions of privacy emphasize individuals' control over definite items of personal information. These conceptions largely do not consider the role of inferential thinking in human thought and data analytics. A gulf may exist between the primary meanings users may intend to convey and the inferential uptake of the information's receiver – indeed, such interpretive space is an important element of how people maintain social relationships and position themselves to different publics. The understandings that are drawn via computer analytics from large, re-contextualized quantities of data and meta-data, however, may be abstract by human standards. Given the capabilities of analytics to generate meanings, limiting disclosure will likely have an increasingly smaller role in individuals' ability to position themselves publicly. Even if we wish greater control over how data winds up in databases, large amounts of data *are* presently being collected and analyzed – and those analytics are a key decisive mechanism in determining who to surveil more closely, who to show which ads, who to target and consider collateral damage for drone strikes, and other aspects of concern to privacy advocates.

This paper recommends that privacy conceptualize disclosure as one part of a chain of meaning-making events. This entails considering in greater depth the meanings we derive from data along with the data itself, and the status it should be afforded in public and personal contexts. Without condoning mass collection and analytic practices, I contend that, for the public sphere, this means asking less of 'what data specifics do we allow into it', and more 'what meanings do we accord precedence'? Privacy offers limited individual protection if it is uncoupled from a wider social context of informational agency. Information is dangerous if it is used to dangerous ends, whether drone bombing or advertising campaign are planned with less or more data.

References

1. Samuel Warren and Louis Brandeis. 1890. The right to privacy. *Harvard Law Review* 4: 193-220.
2. Andrei Marmor. 2015. What is the right to privacy? *Philosophy & Public Affairs.* 43(1): 3-26.
3. Alan Westin. 1967.*Privacy and Freedom.* Simon & Schuster: New York.
4. Alan Westin. 2003. Social and political dimensions of privacy. *Journal of Social Issues.* 59(2).

5. Helen Nissenbaum. 2004. Privacy as contextual integrity. *Washington Law Review* 79: 119-158.
6. Helen Nissenbaum. 2010. *Privacy in context.* Stanford UP: Stanford.
7. Helen Nissenbaum. 2011. A contextual approach to privacy online. *Daedalus.* 140 (4): 32-48.
8. Bert-Jaap Koops et al. 2016. A typology of privacy. *University of Pennsylvania Journal of International Law* 38. Pre-print edition.
9. Ruth Coustick-Deal. 2015. Responding to "Nothing to hide, Nothing to fear." *Open Rights Group* blog. Dec 4. Retrieved 10 March 2016 from https://www.openrightsgroup.org/blog/2015/responding-to-nothing-to-hide-nothing-to-fear
10. Glenn Greenwald. 2014. *No place to hide.* Hamish Hamilton: St. Ives.
11. David Lyon. 2014. Surveillance, Snowden, and Big Data: capacities, consequences, critique. *Big Data & Society* 1: 1 -13.
12. Bruce Schneier. 2015. *Data and Goliath.* Norton and Co: New York.
13. Paul Dourish and Ken Anderson. 2006. Collective information practice: exploring privacy and security as social and cultural phenomena. *Human-Computer Interaction* 21: 319-342.
14. Robert C. Post. 2001. Three concepts of privacy. *Faculty Scholarship Series.* Paper 185. Retrieved 13 May 2016 from http://digitalcommons.law.yale.edu/cgi/viewcontent.cgi?article=1184&context=fss_papers
15. Rob Kitchin. 2013. Big data and human geography: opportunities, challenges and risks. Dialogues in Human Geography 3(3): 262 – 267.
16. Rob Kitchin. 2014. Big Data, new epistemologies, and paradigm shifts. *Big Data and Society* 1.
17. Alice E. Marwick and dannah boyd. 2011. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media and Society.* 13(1): 114-133.
18. Daniel J. Solove. 2007. *The future of reputation.* Yale UP: New Haven.
19. Jon Ronson. 2016. *So you've been publicly shamed.* Macmillan: London.
20. Ferdinand de Saussure. 1916. Cours de linguistique generale. Edited by C. Bally and A. Sechehaye, with the collaboration of A. Riedlinger, Lausanne and Paris: Payot; trans. W. Baskin, *Course in General Linguistics*, Glasgow: Fontana/Collins, 1977.
21. Judith DeCew. 2015. Privacy., *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta, Editor., Retrieved 13 May 2016 from http://plato.stanford.edu/archives/spr2015/entries/privacy/.
22. Daniel Miller et al. 2016. *How the world changed social media.* UCL Press: London. Retrieved 10 March 2016 from http://www.ucl.ac.uk/ucl-press/browse-books/how-world-changed-social-media .
23. Kenneth P. O'Hara et al. Everyday dwelling with WhatsApp. Proceedings of the 17th ACM conference on computer supported cooperative work & social computing. ACM, 2014.
24. Kenneth and Gabrielle Adelman, 2002-2015. About the Streisand lawsuit. Retrieved 16 July 2016 from http://www.californiacoastline.org/ .

25. Daniel Miller. 2016. *Social media in an English village.* UCL Press: London. Retrieved 10 March 2016 from https://www.ucl.ac.uk/ucl-press/browse-books/social-media-in-an-english-village .
26. Geoffrey C. Bowker and Susan Leigh Star. 2000. *Sorting things out.* MIT Press: Cambridge, Massachusetts.
27. David Graeber. 2015. *The utopia of rules.* Melville House: New York.
28. Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2015. Privacy and human behavior in the age of information. *Science.* 30 Jan 2015. 347(6221): 509-514
29. Annemarie Mol. 2008. *The logic of care.* Routledge: New York.
30. Rasika Rampatige, et al. 2014. Hospital cause-of-death statistics: what should we make of them? *Bulletin of the World Health Organization.* Jan 2014. Retrieved 16 July 2016 from https://www.researchgate.net/profile/Lene_Mikkelsen/publications.
31. Mimi Onouhu. 2016. The point of collection. Retrieved 10 March 2016 from https://points.datasociety.net/the-point-of-collection-8ee44ad7c2fa#.jo6he1dvj
32. Don Kulick. 2015. When privacy and secrecy collapse into one another, bad things happen. *Current Anthropology,* 56: Supplement 12: S241-250.
33. Corinne Squire. 2015. Partial secrets. *Current Anthropology,* 56: Supplement 12: S201-10.
34. Joseph Heinrich, Steven J. Heine, and Ara Norenzayan, 2010. The weirdest people in the world? *Behavioral and brain sciences.* 33(2-3), p.61-83.
35. Georg Simmel. 1950. *The Sociology of Georg Simmel*. Wolff K.H. (ed.) The Free Press: New York.
36. Jean-Francois Blanchette. 2011. A material history of bits. *Journal of the American Society for Information Science and Technology*, 62(6): 1042-1057.
37. Anna Reading and Tonya Notley., 2015. The materiality of globital memory: bringing the cloud to earth. *Continuum: Journal of Media & Cultural Studies* 29(4): 511-21.
38. Matthew Kirschenbaum. 2008. *Mechanisms.* MIT Press: Cambridge, Massachusetts, USA.
39. Mirca Madinou and Daniel Miller. 2011. *Migration and new media: transnational families and polymedia.* Routledge: New York.
40. Mirca Madinou and Daniel Miller. 2012. Polymedia. *International Journal of Cultural Studies,* 16,2:169-87.
41. Evan Selinger and Woodrow Hartzog. 2014. Obscurity and privacy. In: *Routledge Companion to Philosophy of Technology.* Joseph Pitt & Ashley Shew, eds.
42. Thomas Nagel. 1998. Concealment and exposure. *Philosophy & Public Affairs* 27(1): 3 - 30.
43. Michael Jackson. 2012. *Between one and one another.* University of California Press: Berkeley.
44. Erving Goffman. 1959. *The presentation of self in everyday life*. Harmondsworth: Penguin.
45. Carla Stang. 2012. *A walk to the river in Amazonia.* Berghahn Books: New York.
46. Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *PNAS.* 110(15): 5802 – 5805.