Discourse analysis and digital surveillance

Discourse Analysis and Digital Surveillance

Rodney H. Jones

Discourse Analysis and Digital Surveillance

Introduction

Some time ago I found a post on my Facebook Newsfeed asking, 'Which Friends are Actually Your Mom and Dad?' (Fig 1). Despite the underlying creepiness of the question, I was curious. So, I clicked on the link and was brought to a dialogue window (Fig 2) asking me if I would like to 'Continue as Rodney?' What it meant was not, did I want to continue being myself, but was I willing to share with a company called 'Meaww World' my 'friends list, timeline posts, and photos'. Undeterred, I clicked 'Okay', and was immediately taken to a page with a pulsing progress bar and the words: '*Calculating result…'* along with an advertisement for a sofa (Fig 3), which was useful given that I had just spent almost an hour browsing online furniture shops. Finally, the answer was revealed on a new page, famed by an array of additional advertisements for furniture, an online supermarket, and other fun quizzes I could take: based presumably on some algorithmic analysis of my personal information, Meaww World had determined that my two friends who could have been my parents were Simon (who could have been my Mom), and Dino (who could have been my Dad). Amused, I immediately posted the result on Facebook, When I did, of course, Simon and Dino, equally amused, shared it on their newsfeeds, and then went on to take the quiz themselves, giving Meaww World access to their data and that of their friends.

Figure 1

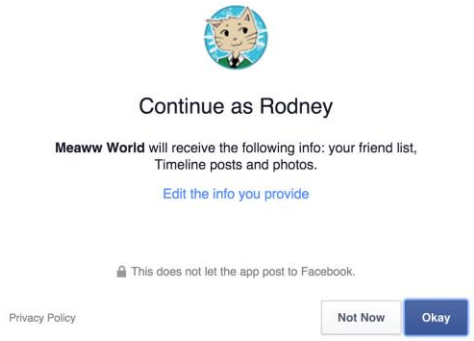
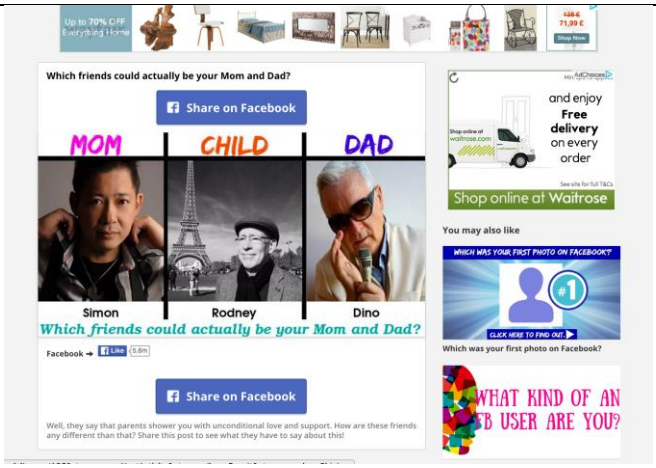


Figure 2



When we think of digital surveillance, we usually think of intelligence gathering by shadowy government agencies of the type uncovered by Edward Snowden. But most digital surveillance is more pedestrian, tied up with our everyday practices of searching the internet, engaging with friends on social media, shopping, and showing off. In fact, as Snowden's revelations chillingly revealed, government surveillance programs like Prism and Mystic are intimately tied to our use of search engines, social media sites, and smartphones, through which we produce vast stores of data that can be exploited by advertisers, government agencies, and

more unsavory actors such as identity thieves and Russian hackers. The ‘Cambridge Analytica scandal’, for example, that may have contributed to the election 2016 of Donald Trump, actually started when researchers convinced Facebook users to take an online quiz called ‘This is Your Digital Life’, through which the company was eventually able to gain access to the data of over 87 million Facebook users (Grassenger & Krogerus, 2017).

This short narrative does not just demonstrate the ubiquitous nature of digital surveillance, it also highlights a number of key-questions about digital surveillance that are relevant to the work of discourse analysts. The first is obviously the role of discourse in making the whole sequence of actions that resulted in me relinquishing my data possible, from the garish scrawl of the original invitation, to the oblique way in which Meaww Wolrd asked for my consent, to the language and layout of the barrage of advertisements that accompanied this process. When we think of digital surveillance, we are accustomed to thinking in terms of complex code, algorithms, and metadata — and, as I will discuss below, these new forms of ‘language’ are indeed central to these processes— but much of what makes digital surveillance possible has to do with the way more traditional ways discourse is deployed to lure users into making compromising decisions. Secondly, there are fundamental questions about ‘speakership’ and ‘listenership’: who exactly is the ‘author’ of the various texts involved here, including the original invitation, the resulting text which I shared with my friends on Facebook, and even the advertisements, which were in part the result of my own internet activities in the hour before I clicked on this quiz? And who is the audience: is it me, Simon and Dino, our various friends who are subjected to this unusual post showing up on their newsfeeds, or Meaww World, who, after all, is the recipient of all of the information we have disclosed? Thirdly, there are questions about ‘meaning’: how did Meaww World come to the conclusion that Simon and Dino could have been

my parents? What was I trying to 'say' by reposting this text on my timeline, and, most importantly, what kinds of meanings are Meaww World and the third parties it might sell my data to able to infer about me based on the data it has gathered? Finally, why did I fall for this scam in the first place? What is it about the discursive strategies of this company, my relationships with my friends on social media, the social practices associated with this medium, and my own experiences of creating and interpreting texts online that made me decide to hand over all of the posts and pictures I have ever sent or received on Facebook to a company that I had never heard of?

In this chapter I will discuss how tools from discourse analysis can contribute to our understanding of digital surveillance, exploring how the interaction among social relationships, discourse practices and technological tools in contemporary digital and physical spaces has created a 'communicative ecology' (Foth & Hearn, 2007) in which nearly all of our social interactions are engineered to produce data of maximal value for internet companies, advertisers and governments. While much of this new communicative ecology is made possible by digital technologies and the sophisticated patterns of participation and methods of discourse processing they make available, much of it also depends on more fundamental practices of human communication that stretch back to the birth of human language itself (Dunbar, 1996), practices like gossip and boasting, and our seemingly insatiable desire to 'see' and 'be seen'.

In what follows, I will first explore what insights from discourse analysis can contribute to our understanding of surveillance more generally. Then I will discuss the *mediated* nature of all surveillance and the different affordances and constraints different media bring to it. In the following section, I will give an overview of the main discursive processes involved in digital surveillance, including participation, pretexting, entextualization, recontextualization, and

inferencing, showing how they occur differently when mediated through digital technologies. Next, I will identify some of the key issues and ongoing debates around digital surveillance related to discourse analysis, specifically identity, agency, and power. I will then go onto discuss the implications of a discourse analytical approach to digital surveillance for the professional practices of applied and sociolinguists. Finally, I will lay out some future directions in which research on discourse and digital surveillance can move.

Information Games

To take a discourse analytical approach to surveillance means taking as a starting point the fact that surveillance is a pervasive fact of everyday life and always has been, even before the advent of social media sites and internet quizzes. Participation in social life is a matter of constantly being watched and 'watching out for others' either to protect them or to protect ourselves from them. As Trottier (2012: 18) observes, 'Surveillance is ubiquitous, not just because of ubiquitous technologies, but because watching and assessing pervade nearly every social relationship.'

This observation is at the heart of much of the work in psychology, sociology and linguistics which informs contemporary approaches to discourse analysis, from Ruesch and Bateson's (1951) observation that the basis of human communication is the 'perception of (being) perceived,' to Goffman's (1964, p. 135) definition of the social situation as 'an environment of mutual monitoring possibilities'. For Goffman, social interaction is essentially a series of ,'information games', involving 'potentially infinite cycle(s) of concealment, discovery, false revelation and rediscovery' (1959, p. 13) through which we negotiate access to various 'territories of the self' (1972), ranging from how visible our bodies and physical actions are to the degree to which we make available the 'preserve' of secrets that we hold inside our heads (or,

nowadays, inside of our digital devices). What is at issue for most people, says Goffman (1972:60), 'is not whether a preserve is exclusively maintained or shared, or given up, but rather (that) the individual is allowed (to determine) what happens to his claim.' What makes surveillance, in the usual sense the word is used, so unnerving for us is the fact that this claim is not being respected, that the information game is not being played 'fairly'. In the word surveillance (Fr. 'looking from above') is an implication of *information asymmetry*, the notion that one party has the upper hand. But for Goffman, the need to address constant potential asymmetries in interaction and to negotiate the balance between what we know about others and what is known about us is an inescapable fact of social life.

All surveillance, then, cannot be regarded as necessarily nefarious or untoward. It is our ability to monitor others, and to negotiate how we are monitored by them, that is to a large degree responsible for our ability to 'take action, seek information and communicate' (Albrechtslund, 2008). Both surveillance and secrecy help produce the social world – they are at the heart of how we interact, manage our relationships, and organize our societies. And, in many situations, such as the one I described above, we willingly make ourselves available for surveillance, perceiving certain psychological or social benefits from being visible to others. 'The negotiation of social identity,' says Phillips (2002, p. 416) 'is not only about the construction and maintenance of boundaries, regions, and performances. It is also about negotiating the permeability of those boundaries.'

Surveillance as Mediated Action

So how do media alter the way these negotiations are accomplished? In order to answer this question, it is good to remember that all surveillance is mediated, that is, it is made possible

by ‘technologies’ of one kind or another, be they windows, keyholes, binoculars, cameras, electronic listening devices, digital media, or just our own eyes and ears, and different media come with different affordances and constraints on who can be surveilled by whom, what kind of information can be gathered, and what can be done with that information (Jones, 2017a).

It is also good to remember that surveillance is never a single action— it always involves a chain of actions, each with explicit goals. Marx (2016), for example lists seven discrete actions that are usually involved in surveillance: 1) tool selection, 2) subject selection, 3) data collection, 4) data processing/analysis, 5) data interpretation, 6) data use, and 7) data fate. Each of these actions may involve the same or different mediational means with the same or different affordances and constraints on how that particular action can be executed.

A range of scholars from sociology, media studies, information sciences and the burgeoning field of surveillance studies have commented on how digital technologies have changed the information game in regard to the different surveillance related actions delineated by Marx. One of the chief ways digital media have changed how surveillance is carried out, for example, is that they make the first two steps almost superfluous — there is no need to make choices about tools and subjects when the tools nearly all subjects are using (such as smartphones and social media accounts) make possible the surveillance of nearly everybody. As Haggerty and Ericson (2000, p. 606) point out, digital media have brought about ‘a rhizomatic levelling of the hierarchy of surveillance, such that groups which were previously exempt from routine surveillance are now increasingly being monitored.’ Just as it is no longer a matter of targeting specific people, it is also no longer a matter of utilizing specific technologies. With the rise of ‘ubiquitous computing’ has come ‘ubiquitous surveillance where wireless sensors are hidden inside of ordinary objects such as cars, kitchen appliances, toilets, buildings and clothes’

(Marx, 2016), and information gathering is increasingly indiscriminate and automated. As Barnard-Willis (2012, p. 22) notes, 'contemporary surveillance is driven by connections between seemingly disparate and previously discrete surveillance technologies, sites, practices and agents,' what Haggerty and Ericson (2000) have famously referred to as 'surveillant assemblages'.

The second important effect of digital technologies scholars have observed has been on the *kinds* of data that can be collected. In 1988 Roger Clarke coined the term 'dataveillance' to describe the 'systematic use of personal data systems in the monitoring of people's actions and communications' (p. 498). Intelligence experts talk about the difference between surveillance engaged in through the actual observation of people (known as 'human intelligence' of 'humint') and that engaged in through the gathering of the 'signals' people give off from their use of technologies (known as 'signals intelligence' or 'sigint'). Digital technologies have made possible a switch in surveillance techniques from the observation of actual persons to the observation of the 'data trails' they leave as they interact with technologies, the incidental 'exhaust' of their mediated actions. This results in an increased focus on information that is, in Goffman's (1963) terms, 'given off', gleaned from actions like clicking, swiping, searching, 'liking' and traveling from one physical location too another, actions that are usually not even regarded as particularly 'meaningful' by those who produce them. In other words, they are not part of the traditional 'preserves' that Goffman says social actors typically guard access to.

What digital technologies make possible is the capture and recording of these manifold incidental actions in digital format, accumulated and aggregated into 'big data' sets that can be operated on by algorithms, which brings me to the third way digital technologies have changed the information game by introducing new and powerful ways of analyzing an interpreting data.

This introduces yet another level of asymmetry between the surveillers and the surveilled: not only are the surveilled unaware what information they produce might be rendered meaningful, but they are also (for the most part) unequipped to interrogate the processes through which this information is interpreted. Tene and Polonetsky (2013:255) compare the relationship between internet platforms and their users to a 'game of poker where one of the players has his hand open and the other keeps his cards close.'

The final way digital technologies have been seen to effect surveillance has to do with the 'use' and ultimate 'fate' of the data collected. While the effects of much digital surveillance, especially of the commercial kind, seem relatively innocuous, resulting in things like targeted advertising and discount offers, in reality the 'social sorting' (Lyon, 2003) that underpins these uses can act to reinforce social and economic inequalities and even lead to dangerous political polarization in the pursuit of profit (Lanier, 2018). Surveillance of any type has the consequences of exasperating the kinds of power asymmetries that made it possible in the first place (Graham, 1998), but the efficiency with which digital technologies are able to 'sort' the subjects of surveillance results in a pervasive prioritization of 'certain people's mobilities, service quality and life chances, while simultaneously reducing those of less favored groups' (such as lower income people and minorities) (Graham & Wood, 2003).

While all of these observations are certainly true, they do not sufficiently capture the *discursive* and *interactive* dimensions of digital surveillance that are so conspicuous in the example with which I began this chapter: the way digital surveillance is made possible both through various kinds of semi-consensual transactions between users and companies and through myriad transactions between users themselves who post, like, share, and otherwise circulate information about one another in a digitized version of Goffman's information game. They also

do not sufficiently highlight how surveillance is crucially supported by complex, chained processes of text production and consumption in which information is continually transformed and recontextualized, its meaning potential changing along the way. Finally, they do not sufficiently address the pragmatic aspects of digital surveillance, the ways technologies are transforming not just how data are 'interpreted', but how meanings are *inferred* more generally. These are aspects of digital surveillance that only an approach informed by discourse analysis can address. While discourse analysis has had some place in surveillance studies, it has mostly been used to critique the way surveillance is represented and justified in texts like newspaper articles and political speeches and through genres such as reality television (e.g. McGrath, 2004; Tiainen, 2017) rather than to examine the actual communicative processes that surveillance entails (Barnard-Wills, 2012). The framework outlined below is intended to highlight those aspects of digital surveillance that are amenable to interrogation by discourse analysts.

Overview: Discourse and Digital Surveillance

A discourse analytical perspective on digital surveillance involves asking what *discursive processes* it entails: what possibilities for interaction, what sorts of interactional roles, and what forms of meaning making. These processes can be divided up into 1) participation: the way digital media create different possibilities for mutual monitoring and different interactional roles and responsibilities; 2) pretexting: the strategies used to compel people to disclose information about themselves; 3) entextualization: the ways digital media facilitate the transformation of actions into text/code and the kinds of meanings that are preserved, lost and changed through these processes; 4) recontextualization: the way information gathered through digital surveillance is transported into different textual contexts and combined with other information to produce

new meanings and new possibilities for action; and 5) inferencing: the ways meanings are inferred from data and used to determine the subsequent kinds of texts and interactions to which users are exposed. Although these processes sometimes happen in succession, as in Marx's (2016) more 'analogue' account of 'surveillance strips', they more often constitute recursive moves in a complex system of feedback loops in which, for example, inferences arrived at through the analysis of data from multiple sources feeds into the formulation of pretexts for the gathering of yet more data. In the following sections I will discuss these processes in more detail.

Participation

The meaning of participation from a discourse analytical perspective has to do not just with who is included in a particular interaction but also what communicative rights and responsibilities different participants have. When it comes to surveillance the key question about participation is: *who is watching whom*? Most models imagine rather static, dyadic participant roles of, as Marx (2012, p. xxv) puts it, a 'surveillance agent (watcher/observer/seeker/inspector/auditor/tester)' and the 'surveillance subject': 'the person about whom information is sought.' Marx does go on, however, to imagine other kinds of participants associated with the 'watcher' role, for example, 'sponsors, data collectors, and initial and secondary users of the data.' Digital security experts have a slightly more elaborate scheme of participants, each with a conventionalized nickname: 'Alice' and 'Bob' for sender and receiver of communications which might be intercepted by 'Eve' (for eavesdropper), 'Carol' (for third person), 'Chuck' (for malicious participant), 'Mallet (for active intruder), 'Trent' (for trusted third party) and 'Grace' (for government agent) (Rivest, Shamir, & Adleman, 1978). But even this more elaborate model of participation is constructed along the binary poles of 'watcher'

and ‘watched’. In networked interactions of the kind described at the beginning of this chapter, however, the distinction between the monitor and the monitored is not so clear-cut. In many ways, in fact, this particular transaction is based upon different actors taking on *multiple* roles as different kinds of ‘watchers’ and different kinds of ‘watched’. One of the main reasons digital media have such a profound impact on the way surveillance is carried out is that they have a profound impact on the way social interaction is carried out more generally, making available more flexible frameworks for participation and creating ways for different kinds of participation frameworks to articulate with one another. As Haggerty (2006, p. 26) puts it, ‘the multiplication of sites of surveillance’ made possible through new technologies

> ruptures the unidirectional nature of the gaze, transforming surveillance from a dynamic microscope to one where knowledge and images of unexpected intensity and assorted distortions cascade from viewer to viewer and across institutions, emerging in unpredictable configurations and combinations, while undermining the neat distinction between watchers and watched through a proliferation of criss-crossing, overlapping and intersecting scrutiny.

This view of participation is, in fact, much more in line with what sociolinguists and discourse analysts, have long noted, that, as Hymes (1974, p. 54) put it, ‘the common dyadic model of speaker-hearer specifies ... too many, sometimes too few, sometimes the wrong participants.’ All social interactions involve the possibilities for multiple modes of production and participation among participants, some ratified, and some unratified (Goffman, 1981). Face-to-face conversations also sometimes have auditors, bystanders, and even eavesdroppers. What technology does is shift the opportunities different participants have for ‘mutual monitoring’.

Media in general contribute to the formation of participation frameworks by enabling and limiting how communication takes place, how it circulates, and who has access to it (Hutchby, 2001). In my previous work (Jones, 2009), I compared the affordances of new media to architectural features like walls and windows which make both surveillance and privacy possible in physical environments. Digital media, I noted, give users tools to negotiate and modulate their visibility. Since then, however, fuelled by advances in mobile technologies and the rise of platforms whose main business model is extracting data from users, digital environments have become more like a house of mirrors (Johnson & Regan, 2014). Negotiations of surveillance and privacy are bound up in a web of multiple tiny acts of people watching each other and offering themselves up to be watched, all driven by algorithmic feedback loops which drive them deeper and deeper into interactions in which they are more likely to disclose more and more information.

These complex new participation frameworks create challenges for people as they try to regulate flows of information and maintain privacy. They also create challenges for discourse analysts, as they attempt to adapt analogue models of speakership and listenership to digital environments (also see chapter by Blommaert, Smits & Yacoubi, this volume). Just as in our analogue interactions, the selves that we construct online are largely a result of ‘a tacit negotiation between ourselves and our imagined auditors (Bowker, 2005, p. 7). On the one hand, people use social media to ‘be seen’, motivated to engage in consensual surveillance by a competition for social status, while on the other, they employ a rage of strategies to avoid ‘over exposure’ such as self-censorship, producing polysemic performances that contain different messages for different audiences (boyd, 2012), using contextualization strategies to negotiate the appropriateness of different messages (Tagg, Seargeant, & Brown, 2017), and alternately

disclosing and withdrawing information in a form of 'virtual identity hide and seek' (Papacharissi & Gibson, 2011, p. 81).

Where older models of participation and audience design are less useful is in understanding the increased role of auditors, bystanders and eavesdroppers in nearly *all* of the interactions we have online. Although Bell (1984) talked about the 'auditor effect' in social interactions, auditors in his model are secondary participants, and eavesdroppers are hardly considered participants at all. For most social media users, however, some awareness of the presence of eavesdroppers is a given, whether they be unratified human participants or faceless corporations mining their data, and a means needs to be developed to measure the degree to which this awareness changes their communication.

Another issue that older models of participation do not address is the presence of what Latour (2007) calls 'new unexpected actors', audiences composed not of humans but of software programs and algorithms who participate in communicative exchanges as '(inter)active co-conspirators' with *both* agents and subjects of surveillance (Hess, 2014). Some people, for example, report altering the way they communicate online based on how they believe algorithms might process and act upon their words, while at the same time, certain interfaces are designed to inspire almost unabashed honesty. As a recent editorial in the IEEE newsletter *Technology & Society* noted: 'We don't lie to our search engine. We're more intimate with it than with our friends' ('Ubiquitous Surveillance and Security', 2017).

Finally, and what is perhaps most striking about the forms of participation enabled by digital media, is that they are almost entirely driven by economic imperatives. On the one hand, internet companies design interfaces that favour certain kinds of interactions over others, offering, for example, the low-cost conviviality of the 'like' button (Jones & Hafner, 2012)

because it has the effect of generating more useful data about people's preferences and personalities. On the other hand, users constantly find themselves engaging in cost-benefit analyses in which the benefits of participation in the network are measured against the costs of exposure of their 'information preserves'. Information has increasingly become the currency we must expend to engage in social life.

Pretextuality

The forms of participation digital media make possible, however, are not enough to make the scale of surveillance internet companies engage in possible, particularly in contexts like the European Union where regulations such as the GDPR require agents of surveillance to make their processes of data collection more transparent (European Comission, 2017). In fact, most surveillance through digital technologies (even much of that conducted by government agents) is carried out within a framework of 'consent' or, at least, 'semi-consent', in which subjects of surveillance are asked to voluntarily relinquish their data. In these circumstances, surveillers must formulate 'pretexts' in order to get people to open their information preserves to scrutiny.

The notion of 'pretexting' comes from the field of social engineering, where it is defined as 'the act of creating an invented scenario to persuade a targeted victim to release information or perform some action' (Hadnagy, 2010, chapter 4). In this context, it is usually associated with con-men or online scammers like 'Nigerian princes' (Blommaert & Omoniyi, 2006). For discourse analysts, the idea of the pretext is much broader. Pretexts are seen as necessary conditions for all communication, sets of expectations text producers and text consumers bring to interaction as a way of negotiating common ground. As Widdowson (2004, p. 79) puts it: 'All texts are designed to be understood pre-textually…it is the pretextual purpose that we bring to

texts that controls how we engage with them.' Another way of understanding pretexts from the discourse analytical perspective is Maryns and Blommaert's (2002, p. 11) definition of them as 'conditions on sayability'—the practices, competencies and contextual frames that make it possible for certain people to credibly engage in certain kinds of interactions. From this perspective, pretexts always involve issues of power as different people bring different pretextual resources to communication. When it comes to pretexts for digital surveillance, both perspectives are relevant: digital pretextuality both creates the contexts for people to disclose personal information and creates what Maryns and Blommaert (p. 11) call 'pretextual gaps', characterised by increasing asymmetries in pretextual resources between surveillers and surveilled.

Internet companies create pretexts for us to surrender information in three ways. First are what might be called 'paradigmatic pretexts' in which scenarios and genres are created to give us reasons to disclose information. These include scenerios like calling a taxi or playing an online game. Palen and Dourish (2003) note that digital media have given rise to a new range of genres which they call 'genres of disclosure', genres like status updates and internet quizzes that are designed to compel users to share information. These genres are embedded in platforms whose 'default settings' are designed to encourage maximum visibility. While users are usually given the opportunity to change default privacy settings, the power of genres and default settings to channel people into certain kinds of behaviour is strong, partially because of the 'transaction costs' involved in changing default settings or resisting the norms associated with particular genres.

The second form of pretexting common in digital environments might be called 'syntagmatic pretexts', those that operate as a result of the ways utterances and actions are sequenced. Widdowson (2004), drawing on the work of Garfinkel (1986), argues that pretexts

are not a matter of static identities and situations; they are about dynamically negotiating identities and situations as 'ongoing accomplishments' through the moment by moment application of 'practical reasoning'. This kind of sequential reasoning comes into play in the design of 'click wraps' and other permission dialogues to gain consent for information gathering which are presented to users either at strategic points in ongoing processes (such as right when users want to take a picture with their phone or locate a place in a map app), so that users are inclined to 'agree' to what's being asked of them, just to get on with the process.

Finally there are what might be called 'emergent pretexts' that arise from interfaces' incessant requests for disclosure that cause users to become 'habituated' to giving consent (Longford, 2005). One example of this is the way the constant requests by websites for permission to use cookies required by laws such as Europe's GDPR has resulted the routinization of consent, an emergent environment in which giving up personal data is seen as a necessary and unremarkable part of accessing information (see below, also Jones 2018, Kim, 2013)

Entextualization

Surveillance has always been about the production and circulation of texts. It is not enough to monitor someone—the results of that monitoring must somehow be 'documented'. As Gitelman (2014) notes, in his treatise on 'paperwork' and modern bureaucracies, the 'knowing-showing' dimension of documentation is not just about recording information, but about controlling the way people understand the social orders that they inhabit. What makes this possible is the process of *entextualization*, the transformation of messages, actions and identities into texts. What is 'known' and 'shown' about subjects of surveillance is largely determined by

how the modes and media employed to document their actions make them ‘legible’ (Scott, 1999).

Bauman and Briggs (1990, p 73) define entextualization as ‘the process of rendering discourse extractable, of making a stretch of linguistic production into a unit-a *text*-that can be lifted out of its interactional setting.’ This process, they say, is always, to some degree, an exercise of power, since those who entextualize reality are able to decide what aspects of reality are captured and how they are represented. They write:

> To decontextualize and recontextualize a text is thus an act of control, and in regard to the differential exercise of such control the issue of social power arises. More specifically, we may recognize differential access to texts, differential legitimacy in claims to and use of texts, differential competence in the use of texts, and differential values attaching to various types of texts. (p. 76)

Digital media have fundamentally changed how entextualization is carried out in practices of surveillance in terms of *what* is entextualized, the *way* it is encoded, and the way these coded artefacts are *circulated*. Whereas in the past, surveillance produced analogue documents and sometimes images which operated within the epistemological boundaries of human language, digital surveillance produces digital documents which operate within a very different epistemology, the logic of code. And what can be ‘known/shown’ through code is very different from what can be ‘known/shown’ through language.

Dodge and Kitchin (2005) make a useful distinction between ‘data’, all possible information about something, and ‘capta’, the information that a system is able to record and store. For computer systems, ‘capta’ consists primarily of streams of gestures — clicks, swipes,

'likes', etc. — that people perform through interfaces -- referred to as 'click-streams'. The mistake many people make when it comes to digital surveillance is being vigilant about what they 'say' online, not about what they *do*: those thousands of tiny actions that produce what Battlle (2005) calls 'a massive click- stream database of desires, needs, wants, and preferences'. The fact is that digital surveillance systems are not yet very good at dealing with the 'content' of our communication. They are much better at recording binary actions in the form of 'metadata'. In the case of surveillance, however, this can actually be an advantage. While too much data can overwhelm surveillance systems and sometimes distort perceptions due to 'noise', more metadata can make analysis more accurate (Blaze, 2013). Collecting metadata is often discounted as less intrusive than actually 'reading people's emails' or 'listening to people's phone calls': as Reilly (2014) puts it, 'metadata is the envelope, not the letter'. Numerous studies (see for example Kosinski, Stillwell, & Graepel, 2013; MIT Media Lab, n.d.), however, have shown that a great deal about people's psychology, personal habits, desires, intentions and even future behaviours can be inferred through metadata. 'Metadata' writes Blaze (2013), 'can reveal far more about us, both individually and as groups, than the words we speak.' As General Michael Hayden, former head of the CIA more bluntly puts it: 'We kill people based on metadata' (Cole, 2014).

Recontextualization

The flip side of entextualization is recontextualization, the ways documents of surveillance are transported and embedded into different contexts, and the ways these different contexts change the meanings of texts and and what they can be used for. For some privacy scholars, the main way surveillance violates privacy is not through the collection of information,

but though the introduction of that information into contexts for which it was not intended, violating what Nissenbaum (2009) calls, ‘contextual integrity’.

A central characteristic of digital media is the fact that they facilitate recontextualzation; they are built upon a logic of ‘data flows’ that encourages social norms of sharing, linking, embedding and re-circulating texts. In daily social interactions, this constant ‘disembedding’ (Giddens, 1991) of actions and utterances from their contexts creates multiple challenges for communication. When it comes to digital surveillance, however, the key aspect of recontextualzation is not the way information is transported into different social contexts as much as it is the way it is transported into different *informational* contexts. The emergence of data management platforms allows surveillers to combine information collected about users not just with information collected about the same users from different sources, but also with information from millions of other users.

Furthermore, this flow of information from dataset to dataset is constant and dynamic. Rather than using the term recontextualization, Haggerty and Ericson (2000) borrow from Deleuze & Guattari (1987) the notions of ‘deterritorialization’ and ‘reterritorialization’, which involve not just inserting information into new contexts but reshaping abstracted content into different forms through combining it with other content. A key aspect of Deleuze & Guattari’s notions of ‘deterritorialization’ and ‘reterritorialization’ is what they call ‘lines of flight’, the different trajectories that information takes and how these trajectories develop their own momentum. With the speed and efficiency of digital networks, it is not just a matter of information being transported from one context to another, but information moving along *trajectories* of recontextualization.

Inferencing

The final way digital surveillance differs from analogue surveillance from a discourse analytical perspective is in how information is *processed* and what is done with it. All surveillance involves a certain amount of 'inferencing' — the data that is gathered is always to some degree incomplete and ultimately has to be 'pieced together' to produce theories about the significance of what subjects of surveillance have said or done. In digital surveillance, these inferences are formed by algorithms working on the kinds of large sets of data described above. But the way algorithms make inferences is very different from the way people do. While humans form inferences through a logic of causation in which we try to discern why someone said what they did by testing it against some mutually agreed upon theory of communication such as Grice's (1989) cooperative principle, algorithms form inferences based on a logic of *correlation* in which 'Meaning is constructed mathematically, probabilistically, based on correlations between pieces of input' (Jones, in press, see also Anderson, 2008; Ayres, 2008). Although these associations are often made based on assumptions about communication and identity held by the people who have written and trained these algorithms, inferences themselves are based on decontextualized Bayesian probabilities rather than on situated human reasoning. This can often result in startlingly accurate inferences, such as when algorithms are able to predict with a high degree of probability a Facebook user's skin colour, political affiliation, sexuality, religion, alcohol, cigarette and drug use, and even whether their parents were divorced on the basis of as few as sixty eight 'likes' (Kosinski et al., 2013), not because of human assumptions about the kinds of things people of a certain ethnicity or political persuasion might 'like', but because of the ability of algorithms to detect patterns in the past 'liking' behaviour of millions of other users. On the other hand, inferences based on big data correlations do sometimes get it wrong,

and when they do, there is no mutually accepted norm of reasoning or theory of communication to appeal to for victims.

Another characteristic of algorithmic inferencing is that it is less about understanding the past as it is predicting the future. In his book on search, John Battelle (2005) calls the stores of information that companies gather about internet users a 'database of intention', an apt phrase because what this database is chiefly used for is making judgments about the probability of future behaviour — what people intend to do next. Moreover, the generation of predictive models based on information gathered from present interactions ends up determining the kinds of interactions that will be available to people in the future. In other words, predictive inferencing ends up creating feedback loops which often end up *causing* the very actions or attitudes that they predict, as, for example, when an algorithm infers that a user is a Donald Trump supporter and so feeds her increasing quantities of pro-Trump content, giving her the impression that pro-Trump sentiments are more widespread than they actually are. Magnet and Gates (2009, p. 3) see this 'cybernetic quality of [digital] surveillance—the capacity to feed personal data about individuals back into the mechanisms of social control' as one of the main things that 'distinguish newer techniques from earlier, less interactive forms of monitoring.'

Key Issues and Ongoing Debates

These discursive changes in the way surveillance is carried out due to the introduction of digital technologies do not just impact the way we understand surveillance, but also force us to rethink a number of key issues in discourse analysis such as identity, agency and power (also see chapter by De Fina & Georgakopoulou, this volume). It has become axiomatic, for example, in discourse analytical circles to regard identity as socially constructed and dynamically negotiated in interaction (see for example Bucholtz & Hall, 2005). The forms of participation,

entextualization and inferencing that support digital surveillance, however, promote a way of understanding identity that assumes that it is 'based upon probabilistic and actuarial logics' (Barnard-Wills, 2012, p. 153) rather than interactional dynamics. This does not mean that the 'data doubles' (Haggerty & Ericson 2000:614) that are constructed of us as a result of digital surveillance are static—like analogue identities, they are also contingent and negotiated — but what they are contingent on is not the social contexts in which we act, but the informational contexts through which our data flows. In his book *We are Data*, Cheney-Lippold (2017) notes how algorithmically produced categories like man, woman, Asian, and wealthy have little to do with the actual social categories of gender, race and class, and more to do with 'a proprietary vocabulary' intended for 'marketers, political campaigns (and) government dragnets.' 'Who we are in the face of algorithmic interpretation' he writes:

> is who we are computationally calculated to be… composed of an almost innumerable collection of interpretive layers, of hundreds of different companies and agencies identifying us in thousands of competing ways. At this very moment, Google may algorithmically think I'm male, whereas digital advertising company Quantcast could say I'm female, and web-analytic firm Alexa might be unsure. (p.5, 6)

Second is the issue of agency and the degree of control we are able to exercise over our actions. In a situation where software allows an increasingly greater degree of mental processes to be delegated to computational systems (Berry, 2011), agency, in the words of Introna (2011, p. 118) 'becomes increasingly encapsulated, nested as codes within codes within codes.' It is not just that many of the information games we play in our social lives are becoming automated, but also that software might be 'training' us to surrender our agency. A good example of this is the issue of consent, where the form and volume of dialogue boxes asking us to grant consent give

rise to what Rock calls ‘tick box consent’ (2016). Consent gathering processes are often engineered ‘to skew individual decision-making, in effect creating an illusion of free choice that helps to legitimatize surveillance practices’ (Kerr, Lucock, & Steeves, 2009, p.). Unfortunately, most technical and legislative solutions which aim to help people regain control of their data are based on the flawed assumption that the more we are asked to give our consent for data gathering, the more control we have over it. On the contrary, the deluge of permission dialogues that are required by such laws as Europe’s General Data Protection Regulations likely results in a kind of ‘consent fatigue’, with users more and more likely to suspend judgement and just click ‘agree’ (Jones, 2018).

Perhaps the most important tasks of discourse analysts, however, is understanding how electronic surveillance changes the way power is discursively constituted and exercised. What has been called the ‘information revolution’, says Andrejevic (2015, p. x), is better conceptualized as a surveillance revolution in which exercises of power are increasingly ‘informatic’. This power operates on both the maco-level, as large platforms hoard vast stores of data which allow them to dictate the parameters of communication, commerce, and, increasingly, governance, and, on the micro-level, as the surveillance agendas of the platforms dictate the kinds of social interactions and the kinds of life chances individuals can have. What has been missing from critiques of power in digital surveillance is a thorough understanding of how power is enabled and exercised through the discursive dimensions of surveillance, how, for example, the ‘encoding of human experience’ forces users of digital media to ‘submit to particular ways of categorizing and conceptualizing the world’ (Duranti, 2011, p. 29), how asymmetries in information processing capabilities create ‘pretextual gaps’ (Maryns & Blommaert, 2002), and how, through the transformation of bodies into information, individuals are fragmented,

disciplined and denied authentic opportunities for communication in which identities are situated, relational and mutually negotiated (Anthamatten, 2015).

Implications for professional practice

Whether we like it or not, linguists are deeply implicated in the creation and maintenance of the modern surveillance apparatus. There is, for example, a long history of linguists cooperating with intelligence services in areas like language teaching and cryptology. Linguists contribute to the monitoring and disciplining of migrants through lending their forensic expertise to the evaluation of claims for refugee status and developing language tests which operate as tools of normalization and control (Shohamy, 2014). Linguists write the natural language processing algorithms used to capture and analyse both intelligence data and data on consumer behaviour, and even in our teaching we submit our students to a wide array of surveillance practices, many enabled by digital media such as plagiarism detection software and learning analytics platforms.

At the same time, applied linguists and discourse analyst can also be part of the solution, contributing to a robust critique of digital surveillance and helping citizens to regain agency and control over their information. Discourse analysts, for example, can help to formulate design solutions such as interfaces that more effectively elicit genuine consent from users when their data is being collected or enable them to monitor the data they are giving to others (Nguyen & Mynatt, 2002). They can also help to design and implement curricula that alert students to the discursive aspects of digital surveillance and give them opportunities to productively reflect upon how they interact with both human and non-human actors within surveillant assemblages (see for example Jones, 2017b). Finally, applied linguists and discourse analysts can make a contribution to advocacy and critique, both in the workplace by, for example raising ethical issues about the

use of digital platforms to gather data about students and staff (Slade & Prinsloo, 2013), and outside the classroom, calling attention to policies that allow governments and private entities to gather data about citizens and helping to document and critique how digital surveillance affects the experiences and rights of migrants and refugees (see for example Khan, 2019).

Future Directions

The main aim of this chapter has been to make a case for the application of tools from discourse analysis to understanding and critiquing digital surveillance. Future work in this area will need to respond both to rapid technological changes which introduce increasingly sophisticated ways to gather information about people and theoretical and methodological advances in the field of discourse analysis.

One key-development that will need to be addressed is the rise of biometric technologies, which present challenges to the way we understand how bodies and information are interconnected (Ploeg, 2003), challenges which dovetail with trends in sociolinguistics that advocate for more attention to the embodied dimensions of communication (see for example Bucholtz & Hall, 2016). As Hayles (1993, p. 162) writes: 'embodiment mediates between technology and discourse by creating new experiential frameworks that serve as boundary markers for the creation of corresponding discursive systems.'

Another important challenge will be the increasing sophistication of intelligent agents embedded in all sorts of mundane objects like toothbrushes, television sets and automobiles. Here the recent turn in applied linguistics toward post-human theory (Pennycook, 2017) holds the promise of helping us to understand how the 'agency' of objects affects how we play the information game (also see chapter by Lamb & Higgins, this volume). Increased interest in the

role of 'affect' in digital communication (Bucher, 2017, Wee, 2015) will also contribute to future research on the discursive and interactive dimensions of digital surveillance, addressing how the new 'affective economies' (Ahmed, 2004) online might change how people communicate (also see chapter by Giaxoglou & Seargeant, this volume).

The most important thing for discourse analysts tackling the issue of digital surveillance will be to avoid beginning with *a priori* assumptions about the roles and identities of 'surveillers' and 'surveilled' and to focus on the complex and contingent nature of online surveillance in which both humans and non-human actors assume multiple positions, construct multiple identities and engage in myriad micro-strategies of compliance or resistance moment by moment though the medium of discourse (Barnard-Wills, 2012).

Summary

This chapter has presented an overview of the discursive dimensions of digital surveillance. It began with an account of the ways other disciplines have seen the relationship between surveillance and digital media. It then went on to argue that surveillance is, to some degree, part of all interactions as individuals and institutions employ different discursive strategies to conceal or disclose information about themselves and gather and process information about others. The chapter explained five aspects of digital surveillance relevant to discourse analysts: 1) participation (the way digital media make available unique opportunities for 'mutual monitoring'), 2) pretexting (the strategies used to get users of digital media to surrender their data); 3) entextualization (the way digital media facilitate the capture and encoding of data); 4) recontextualization (the way digital media affect the way data are circulated); and 5) inferencing (the way inferences are made based on datasets). The chapter then went on to identify some key

issues in discourse analysis impacted by digital surveillance as well as some implications for the professional practices of discourse analysts and other scholars of language. The chapter ended by considering how more recent trends in discourse analysis around such notions as embodiment (see chapter by Busch, this volume), post-humanism, and affect might inform future work on discourse and digital surveillance.

Further Reading

Barnard-Wills, D. (2012). *Surveillance and Identity: Discourse, Subjectivity and the State*. Farnham: Ashgate.

This is one of the few examples in surveillance studies to seriously engage with theories of discourse analysis, exploring how regimes of surveillance discursively construct social identities and relationships between citizens/customers and the state and corporations.

Jones, R. (2017) Surveillant media: Technology, language and control. In C. Cotter and D. Perrin (eds.) *The Routledge Handbook of Language and Media*. London, Routledge, 244-261.

This chapter gives a comprehensive explanation of the impact of different kinds of media on the discursive processes involved in surveillance.

Jones, R. (accepted); 'Folk algorithmics': Reading and writing in the age of the algorithm. *Linguistics and Education*.

This article looks at digital surveillance from the point of view of *users* of digital media, reporting on a research project in which participants described their 'folk theories' of how algorithms work and reflected on how these theories affected the way they communicated.

References

Ahmed, S. (2004). Affective Economies. *Social Text*, *22*(2), 117–139.

Albrechtslund, A. (2008). Online Social Networking as Participatory Surveillance. *First Monday*, *13*(3).

Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired Magazine,* June 23. Retrieved 19 April 2018, from https://www.wired.com/2008/06/pb-theory/

Andrejevic, M. (2015). Forward. In *Feminist Surveillance Studies* (pp. ix–xviii). Durham, NC: Duke University Press.

Ayres, I. (2008). *Super Crunchers: How Anything Can Be Predicted*. London: Hachette UK.

Anthamatten, E. (2015, March 23). Visibility Is a Trap: Body Cameras and the Panopticon of Police Power. Retrieved 6 September 2018, from http://www.mantlethought.org/philosophy/visibility-trap

Barnard-Wills, D. (2012). *Surveillance and Identity: Discourse, Subjectivity and the State*. Farnham: Ashgate.

Battelle, J. (2005). *The Search: How Google and its Rivals Rewrote the Rules of Business and Transformed our Culture*. Portfolio.

Bauman, R., & Briggs, C. L. (1990). Poetics and performance as critical perspectives on language and social life. *Annual Review of Anthropology*, *19*, 59–88.

Bell, A. (1984). Language style as audience design. *Langauge in Society*, *13*, 145–204.

Berry, D. (2011). *The Philosophy of Software: Code and Mediation in the Digital Age*. Houndmills, Basingstoke, Hampshire : New York: Palgrave.

Blaze, M. (2013, June 19). Phew, NSA Is Just Collecting Metadata. (You Should Still Worry). *Wired*. Retrieved from https://www.wired.com/2013/06/phew-it-was-just-metadata-not-think-again/

Blommaert, J., & Omoniyi, T. (2006). Email fraud: language, technology, and the indexicals of globalisation. *Social Semiotics*, *16*(4), 573–605.

Bowker, G. C. (2005). *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.

boyd, D. (2012). Networked Privacy. *Surveillance & Society*, *10*(3/4), 348–350.

Bucholtz, M., & Hall, K. (2005). Identity and interaction: a sociocultural linguistic approach. *Discourse Studies*, *7*(4–5), 585–614. https://doi.org/10.1177/1461445605054407

Bucholtz, M., & Hall, K. (2016). Embodied sociolinguistics. In N. Coupland (Ed.), *Sociolinguistics* (pp. 173–198). Cambridge: Cambridge University Press.

Cheney-Lippold, J. (2017). *We Are Data: Algorithms and The Making of Our Digital Selves*. New York: NYU Press.

Clarke, R. (1988). Information Technology and Dataveillance. *Commun. ACM*, *31*(5), 498–512.

Cole, D. (2014, May 10). ‘We Kill People Based on Metadata’. Retrieved 4 September 2018, from https://www.nybooks.com/daily/2014/05/10/we-kill-people-based-metadata/

Deleuze, G., Guattari, F., & Massumi, B. (1987). *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: University of Minnesota Press.

Dodge, M., & Kitchin, R. (2005). Codes of Life: Identification Codes and the Machine-Readable World. *Environment and Planning D: Society and Space*, *23*(6), 851–881.

Dunbar, R. (1996). *Grooming, Gossip and the Evolution of Language*. Cambridge, MA: Harvard Univerity Press.

Duranti, A. (2011). Linguistic anthropology: The study of language as a non-neutral medium. In R. Mesthrie (Ed.), *The Cambridge Handbook of Sociolinguistics* (pp. 28–46). Cambridge: Cambridge University Press.

European Comission. (2017). 2018 reform of EU data protection rules. Retrieved 2 September 2018, from https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

Foth, M., & Hearn, G. (2007). Networked Individualism of Urban Residents: Discovering the communicative ecology in inner-city apartment buildings. *Information, Communication & Society*, *10*(5), 749–772.

Garfinkel, H. (1986). Remarks on ethnomethodology. In J. J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics: The Ethnography of Communication* (2nd ed., pp. 301–345). New York: John Wiley & Sons.

Giddens, A. (1991). *The Consequences of Modernity*. Cambridge, UK: Polity Press.

Gitelman, L. (2014). *Paper Knowledge: Toward a Media History of Documents*. Durham, NC: Duke University Press.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York: Doubleday.

Goffman, E. (1963). *Stigma: Notes on the Management of Spoiled Identity*. Englewood Cliffs, NJ: Prentice Hall.

Goffman, E. (1964). The neglected aituation. *American Anthropologist*, *66*(6_PART2), 133–136.

Goffman, E. (1972). *Relations in Public: Microstudies of the Public Order*. Harper & Row.

Goffman, E. (1981). *Forms of talk*. Oxford: Blackwell.

Graham, S. (1998). Spaces of surveillant simulation: New technologies, digital representations, and material geographies. *Environment and Planning D: Society and Space*, *16*(4), 483–

Graham, S., & Wood, D. (2003). Digitizing surveillance: Categorization, space, inequality. *Critical Social Policy*, *23*(2), 227–248.

Grassenger, H., & Krogerus, M. (2017, January 28). The data that turned the world upside down. Retrieved 18 August 2018, from https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Hadnagy, C. (2010). *Social Engineering: The Art of Human Hacking*. Indianapolis, IN: Wiley.

Haggerty, K. D. (2006). Tear down the walls: on demolishing the panopticon. In D. Lyon (Ed.), *Theorizing Surveillance* (pp. 23–45). London: Routledge.

Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *British Journal of Sociology*, *51*(4), 605–622.

Hayles, N. K. (1993). The materiality of informatics. *Configurations*, *1*(1), 147–170.

Hess, A. (2014). You are what you compute (and what is computed for you): Considerations of digital rhetorical identification. *Journal of Contemporary Rhetoric*, *4*(1/2), 1–18.

Hutchby, I. (2001). *Conversation and Technology: From the Telephone to the Internet*. Cambridge, UK ; Malden, MA: Polity.

Hymes, D. (1974). *Foundations in Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press.

Introna, L. D. (2011). The enframing of code agency, originality and the plagiarist. *Theory, Culture & Society*, *28*(6), 113–141.

Johnson, D. G., & Regan, P. M. (2014). *Transparency and Surveillance as Sociotechnical Accountability: A House of Mirrors*. London: Routledge.

Jones, R. H. (2009). Inter-activity: How new media can help us understand old media. In C. Rowe & E. Wyss (Eds.), *New Media and Linguistic Change* (pp. 11–29). Cresskill, NJ: Hampton Press.

Jones, R. H. (2017a). Surveillant media: Technology, language and control. In C. Colleen & D. Perrin (Eds.), *The Routledge Handbook of Language and Media* (pp. 244–261). London: Routledge.

Jones, R. H. (2017b). 'The text is reading you': Language teaching in the age of the algorithm. Presented at the 18th World Congress of Applied Linguistics, Rio de Janeiro.

Jones, R. H. (2018). GDPR and the discursive coercion of consent. A plenary address presented at the 'WhatsUp Switzerland' conference: Language, Individuals and Ideologies in Mobile Messaging, University of Zurich, October 18-20.

Jones, R. H. (in press). The rise of the pragmatic web: Implications for rethinking meaning and interaction. In C. Tagg & M. Evans (Eds.), *Historicising the digital*. Amsterdam: De Gruyter Mouton.

Jones, R. H., & Hafner, C. A. (2012). *Understanding Digital Literacies: A Practical Introduction*. London: Routledge.

Kerr, I., Lucock, C., & Steeves, V. (2009). *Lessons from the Identity Trail: Anonymity, Privacy and Identity in a Networked Society* (1 edition). Oxford ; New York: Oxford University Press.

Khan, K. (2019). *Becoming a Citizen: Linguistic Trials and Negotiations in the UK*. London, UK ; New York, NY: Bloomsbury Academic.

Kim, N. S. (2013). *Wrap Contracts: Foundations and Ramifications*. Oxford: Oxford University Press.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802–5805.

Lanier, J. (2018). *Ten Arguments for Deleting your Social Media Accounts Right Now*. New York: Henry Holt and Company.

Latour, B. (2007). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford; New York: Oxford University Press.

Longford, G. (2005). Pedagogies of digital citizenship and the politics of code. *Techné: Research in Philosophy and Technology*, *9*(1).

Lyon, D. (2003). *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*. New York: Psychology Press.

Magnet, S., & Gates, K. (2009). Communicating surveillance: Examining the intersections. In K. Gates & S. Magnet (Eds.), *The New Media of Surveillance*. London: Routledge.

Marx, G. (2012). “Your papers please”: personal and professional encounters with surveillance. In D. Lyon, K. Ball, & K. D. Haggerty (Eds.), *Routledge Handbook of Surveillance studies* (pp. xx–xxxi). London: Routledge.

Marx, G. T. (2016). *Windows into the Soul: Surveillance and Society in an Age of High Technology*. Chicago ; London: University of Chicago Press.

Maryns, K., & Blommaert, J. (2002). Pretextuality and pretextual gaps: on re/defining linguistic inequality. *Journal of Pragmatics*, *12*(1), 11–30.

McGrath, J. (2004). *Loving Big Brother: Surveillance Culture and Performance Space*. London ; New York: Routledge.

MIT Media Lab. (n.d.). Project overview: On the reidentifiability of credit card metadata. Retrieved 6 September 2018, from https://www.media.mit.edu/projects/on-the-reidentifiability-of-credit-card-metadata/overview/

Nguyen, D. H., & Mynatt, E. D. (2002). *Privacy Mirrors: Understanding and Shaping Socio-Technical Ubiquitous Computing Systems* (Technical Report). Georgia Institute of Technology. Retrieved from https://smartech.gatech.edu/handle/1853/3268

Nissenbaum, H. (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Palo Alto: Stanford University Press.

Palen, L., & Dourish, P. (2003). Unpacking 'privacy' for a networked world. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 129–136). New York, NY, USA: ACM.

Papacharissi, Z., & Gibson, P. L. (2011). Fifteen minutes of privacy: Privacy, sociality, and publicity on social network sites. In S. Trepte & L. Reinecke (Eds.), *Privacy Online: Perspectives on Privacy and Self-Disclosure in the social web* (pp. 75–89). Berlin, Heidelberg: Springer Berlin Heidelberg.

Pennycook, A. (2017). *Posthumanist Applied Linguistics*. London: Routledge.

Phillips, D. J. (2002). Negotiating the digital closet: Online pseudonymity and the politics of sexual identity. *Information, Communication & Society*, *5*(3), 406–424.

Ploeg, I. van der. (2003). Biometrics and privacy A note on the politics of theorizing technology. *Information, Communication & Society*, *6*(1), 85–104. https://doi.org/10.1080/1369118032000068741

Reilly, C. (2014, August 13). The metadata debate: What you need to know about data retention. Retrieved 4 September 2018, from https://www.cnet.com/news/what-you-need-to-know-about-data-retention/

Rivest, R. L., Shamir, A., & Adleman, L. (1978). A method for obtaining digital signatures and Public-key Cryptosystems. *Commun. ACM*, *21*(2), 120–126.

Rock, F. (2016). Talking the ethical turn: Drawing on tick-box consent in policing. In S. Ehrlich, D. Eades, & J. Ainsworth (Eds.), *Discursive Constructions of Consent in the Legal Process* (pp. 93–117). Oxford: Oxford University Press.

Ruesch, J., & Bateson, G. (1951). *Communication: The Social Matrix of Psychiatry*. New York: W.W. Norton & Cpmpany.

Scott, J. (1999). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press.

Shohamy, E. (2014). *The Power of Tests: A Critical Perspective on the Uses of Language Tests*. Routledge.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529.

Tagg, C., Seargeant, P., & Brown, A. A. (2017). *Taking Offence on Social Media: Conviviality and Communication on Facebook*. New York: Palgrave Macmillan.

Tene, O., & Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property*, *11*(5), 239.

Tiainen, M. (2017). (De)legitimating electronic surveillance: a critical discourse analysis of the Finnish news coverage of the Edward Snowden revelations. *Critical Discourse Studies*, *14*(4), 402–419.

Trottier, D. (2012). *Social Media as Surveillance: Rethinking Visibility in a Converging World*. Surrey: Ashgate.

Ubiquitous Surveillance and Security. (2017, June 29). Retrieved 22 August 2018, from http://technologyandsociety.org/ubiquitous-surveillance-and-security/

Wee, L. (2015). Mobilizing affect in the linguistic cyberlandscape: The R-Word Campaign. In R. Rudby & S. Ben Said (Eds.), *Conflict, exclusion and dissent in the linguistic landscape* (pp. 185–203). Basingstoke: Palgrave Macmillan.

Widdowson, H. G. (2004). *Text, Context, Pretext: Critical Issues in Discourse Analysis*. Wiley.

.

Bio Note

Rodney H. Jones is Professor of Sociolinguistics and Head of the Department of English Language and Applied Linguistics at the University of Reading. His recent books include *Discourse and Digital Practices* (Routledge, 2015), *Spoken Discourse* (Bloomsbury, 2016) and *A Sociolinguistics of Surveillance* (Oxford, forthcoming).