

“It’s Not Like It’s Life or Death or Whatever”: Young People’s Understandings of Social Media Data

Social Media + Society
July–September 2018: 1–9
© The Author(s) 2018
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2056305118787808
journals.sagepub.com/home/sms

Luci Pangrazio¹ and Neil Selwyn²

Abstract

Young people’s engagements with social media now generate large quantities of personal data, with “big social data” becoming an increasingly important “currency” in the digital economy. While using social media platforms is ostensibly “free,” users nevertheless “pay” for these services through their personal data—enabling advertisers, content developers, and other third parties to profile, predict, and position individuals. Such developments have prompted calls for social media users to adopt more informed and critical stances toward how and why *their* data are being used—that is, to build “critical data literacies.” This article reports on research that explores young social media users’ understandings of their personal data and its attendant issues. Drawing on research with groups of young people (aged 13–17 years), the article investigates the consequences of making third party (re)uses of personal data openly available for social media users to interpret and make critical sense of. The findings provide valuable insights into young people’s understandings of the technical, social, and cultural issues that underpin their ability to engage with, and make sense of, social media data. The article concludes by considering how research into critical data literacies might connect in more meaningful and effective ways with everyday lived experiences of social media use.

Keywords

social media, personal data, data literacies, teenagers, privacy, security

Introduction

The mass use of social media platforms has given rise to unprecedented rates of data generation and (re)circulation. Implicit to social media participation is the sharing of personal information in order to communicate and interact with “friends” and “followers” online. Yet, the conscious volunteering of user information and “profile” details only represents a fraction of the personal data generated, as individuals also generate data *unconsciously*. Everything from platform interactions to the content of posts, photos, videos, and associated technical details can be used as “social media data” (Kennedy & Moss, 2015), which relate specifically to individual users. Due to the opacity of the digital infrastructure, individuals often have little insight and understanding of the ways in which these data might be used in the future. Of course, individual user data are an integral element of the technical operation of social media platforms. Yet as many previous articles in this journal have highlighted, these data are also integral to the commercial viability of social media (Helmond, 2015; Papathanassoupoulos, 2015; Shade & Singh, 2016). Indeed, social media platforms that are

ostensibly “free” of charge are monetized around the selling of “big social data” to third parties—often commercial and government organizations looking to better profile consumers, target advertising, and refine products and services.

The symbiotic relationship between social media users and their data is now the focus of considerable academic interest. These social data are an important part of how individuals’ online engagements and experiences are shaped—from the advertising that each individual will encounter through to the services that one is permitted to access. Moreover, these data are an increasingly significant part of how the actions of individuals are enabled and/or constrained by dominant institutions and organizations in their lives. Big

¹Faculty of Arts & Education, Deakin University, Australia

²Faculty of Education, Monash University, Australia

Corresponding Author:

Luci Pangrazio, Faculty of Arts & Education, Deakin University, 221 Burwood Highway, Melbourne, VIC, Australia, 3125.
Email: luci.pangrazio@deakin.edu.au



social data are an important component of decision-making in fields ranging from financial credit through to job recruitment. Concerns have therefore been raised over the transparency and fairness of these processes, particularly in terms of the ability of social media users to engage with these data processes in informed and empowered ways. Depictions of “the used user” (Peacock, 2014) reflect growing academic concerns over “the asymmetrical power relations that are deeply imbricated in the structural ways in which data are produced by and yet flows away from the user” (Pybus, Cote, & Blanke, 2015, p. 4).

Thus, it is now being argued that individuals need to develop informed and critical stances toward how and why *their* data are being used. This is often framed in terms of individuals building “critical digital literacies” and becoming vigilant “data citizens” (Gregory & Bowker, 2014). These issues are seen to be especially pertinent for teenage social media users (i.e., aged 12–18 years). This age group constitutes some of the most voracious but vulnerable users of social media. On one hand, it is widely accepted that “smart” management of personal data can enhance young people’s use of digital technology. However, recent European and North American research has raised concerns over a sense of powerlessness among teenagers with regard to personal data, and their limited control over data privacy and security (Donovan, 2013; Pybus et al., 2015). To this end, enhanced awareness and control of personal data is now being seen as a crucial part of supporting young people’s engagement with social media.

Yet, this is not a straightforward issue that can be addressed easily. First, social media platforms are designed and configured in ways that compel many users to simply acquiesce to the continuous sharing of data. As Custers (2016) observes, “browsing and surfing would take a lot of time if every internet user would really think through every consent request that is asked for” (p. 3). Second, there are clear limitations to technical and legislative controls. For example, while most social media platforms rely on user agreements to “Terms Of Service” and privacy policies, “there is growing skepticism regarding the effectiveness of informed consent in the context of personal data processing” (Custers, 2016, p. 2). Similarly, technical fixes such as encryption software, data-blocking software, and privacy filters are seen to require technical skills and competencies beyond those of many technology consumers (young and old alike) (Matzner, Masur, Ochs, & von Pape, 2016).

Third, it also appears that personal data are an aspect of social media that many users are unwilling and/or unable to approach in a critical manner. While data privacy is found to be a matter of considerable concern for young people (e.g., Marwick & Boyd, 2014), this tends to relate predominantly to the active sharing of content to others (Suh & Hargittai, 2015). Besides concerns over passing content to “friends” and “friends of friends,” third party (re)appropriation of personal data is often *not* seen as an immediate issue by social

media users. Indeed, Suh and Hargittai (2015) report that many young people see more value in being active and sharing information on Facebook than protecting their personal information, meaning data-mining is seen as an acceptable tradeoff for platform participation.

While it might not be a topic of particular interest to most social media users, there is a growing academic concern with addressing the “intellectual detachment” (Obar, 2015) of individuals with regard to their personal data. This issue has certainly come to the fore in the emergent field of “critical data studies” (e.g., Dalton et al., 2016; Kitchin & Lauriault, 2018). This approach highlights the need for better understandings of how digital data are implicated in shaping what people can and cannot do in the shaping of opportunities and—in short—in the operation of power.

In these terms, researchers have begun recently to consider the politics of young people’s personal digital data. In particular, a few studies have explored ways of supporting young people’s “data agency”—what Pybus et al. (2015) describe as “a recursive data public . . . with augmented critical data making capacity” (p. 8). This is seen to involve supporting young people to become “fully informed agents” (Peacock, 2014, p. 7) whose social media practices are grounded in sophisticated understandings of how personal data are generated and (re)used as a result of their actions. As Peacock (2014) continues,

As a working thesis, I propose that online users do not use the internet to “donate” personal data to unknown corporate entities A case can be made [. . .] for an inherent preference amongst individuals to control the extraction and distribution of personal data. People prefer to be the experts of their own situation. (p. 2)

Research Questions and Methods

This article therefore responds to ongoing calls for research that seeks to explore “how social media data and their mining might be made accessible to publics” (Kennedy & Moss, 2015, p. 7). Specifically, the remainder of the article reports on research designed to explore the extent to which young social media users are aware of their personal data and its attendant issues. As such, the article addresses the following research questions:

- What are young people’s understandings of personal data, and how do these relate to their social media practices?
- What concerns do young people express about personal data associated with their social media use?
- What new personal data practices do young people see as appropriate and/or attainable?

These questions were addressed by a 12-month research study, which sought to work with groups of young people to cultivate “critical” understandings of their own personal

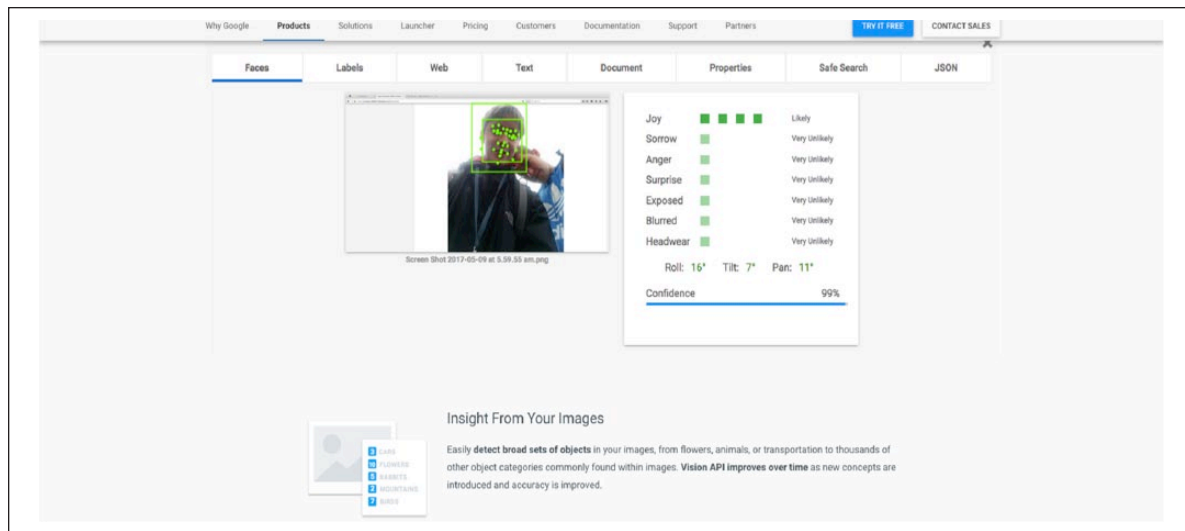


Figure 1. Sample sentiment analysis from PDQ dashboard link.

data. The project ran from July 2016 through June 2017, working with five groups (overall $n=27$) in the Melbourne metropolitan area of Australia, composed of 11 females, 14 males, and 1 non-binary participant aged 13–17 years.

The project adopted methods that are beginning to be used in social and computational science to make data processing more visible and accessible to participants (Kennedy & Moss, 2015; Pybus et al., 2015). This involved a combination of code-based experimentation with digital media and data generation (Haber, 2016), as well as a participatory workshops (Blomberg & Henderson, 1990; Bødker, Grønbaek, & Kyng, 1993) designed to support participants to co-construct alternative ways of engaging with social media along “data savvy” lines. The workshops were used as a means to investigate the realities of young people’s social media use and to highlight where it might be possible to reconfigure social media practices more democratically in light of how personal data are being re-used and re-appropriated by third parties (see Ehn, 1988). As such, participatory design offered an ideal means of testing the potential and limitations of “critical” personal data principles against the confines of the young people’s everyday lives.

Following this approach, the first phase of investigation involved establishing the young people’s current understandings of personal data. The second stage involved generating a sample of personal data, which participants could analyze and reflect upon. In order to facilitate this, the researchers first worked with a software developer to develop an Android smartphone chat app (titled “PDQ”) that was capable of aggregating individuals’ personal data, and then demonstrating to each participant how their data might be recirculated and reused by various third parties. Participants who did not have an Android phone were supplied with one for the duration of the study. At the conclusion of the study, the app and data were deleted from all of the mobile phones. The app was

developed to provide insights into three forms of data generated by social media use:

1. The use of natural language processing to extract metadata from chat logs (using the Cloud Language API);
2. The use of shared images to extract metadata on the user, including sentiment analysis of faces and object recognition (using the Cloud Vision API);
3. The use of GPS co-ordinates from the users’ Android device to map user whereabouts and reverse geo coding to create user location meta-data (using Google Maps API).

The project’s use of a bespoke mobile app was deliberate. First, privacy issues regarding young people’s social media use are seen to increase with the use of mobile devices (Suh & Hargittai, 2015). Second, the app’s use of industry-practice APIs was intended to open up access to commercial data mining techniques, and tools that are normally hidden—and therefore rendered uncontested—to ordinary users (Kennedy & Moss, 2015). As Boyd and Crawford (2012) put it, “wrangling APIs, scraping, and analyzing big swathes of data is a skill set generally restricted to those with a computational background” (p. 674). In a practical sense, PDQ generated insights into the sentiments of individuals depicted in images (see Figure 1); the locations individuals visited over the course of the week, presented as a list of commercial outlets (i.e., McDonalds, shops), institutions (i.e., schools, doctors surgeries), and community venues (i.e., local swimming pools); and analysis of text for references to events, documents, and people. The app is currently undergoing further development so that it will be more broadly available for research and educational purposes.

Once participants had used PDQ to generate a sample of personal data over a 7-day period, they then engaged in a series of three “personal data workshops” that focused on: (1) understanding personal data practices, (2) linking these with personal data trails and traces, and (3) identifying methods of personal data protection and/or resistance. This final workshop used participatory design methods to help young people devise strategies and tactics that could help them to better manage and make sense of their personal data. At each of the three phases of the investigation, participants were split up into groups of 3–6 for focus group discussions conducted by the researchers. Each workshop therefore generated large amounts of data relating to the young people’s developing understandings of their personal data and social media use.

Researchers systematically collected data from all the workshops in the form of observations and recorded discussions. Data were analyzed in three sweeps, including (1) an initial reading involving memoing, selecting, summarizing and coding; (2) subsequent stages of construction of themes guided by the research questions and background theory; and (3) final stages involving finer-grained coding involving constant comparison and theory building (Denzin & Lincoln, 2005; Miles & Huberman, 1984). Categories of analysis were developed which related both to the background theory and to the analysis of the data. What follows is an analysis of these data with regard to the three research questions.

Findings

These young people could be described as eclectic users of social media, with varied uses for the different platforms they engaged with. The most popular social media platforms used were Snapchat and Instagram, with one participant describing Instagram as “the app of a generation.” By contrast, very few made use of Facebook, as it was seen as a social media platform for adults, or the “mum social media thing” [Zara]. In fact, many participants appeared wary of Facebook and described it as “very public” [Mac] and “complicated” [Hannah]. This was in contrast with the other platforms they used which were perceived as more secure and simpler to use. Several participants eschewed what they described as “mainstream social media platforms” [Blair] and predictable patterns of use, and instead used social media for specific purposes. For example, Andrew used Twitter for the sole purpose of locating celebrities when they were in town. While three girls appropriated Snapchat by covering the camera on their phone to send “snaps” *without* images, resisting conventional or expected uses of the platform. As such, this group could be characterized as using social media in ways that were responsive to the cliques and niches of young people.

Young People’s Understandings of Personal Data at the Beginning of the Workshops

At the beginning of the workshop sessions, participants were interviewed about their understandings of personal data.

These interviews highlighted several areas of concern and awareness, as well as other issues that were less of a concern or non-concerns.

Concerns Pre-PDQ. The young people’s concerns over personal data centered predominantly on (1) “intimate” individuals that they had strong interpersonal ties with; (2) technology and telecommunications companies, whom they regarded as tracking and recording their personal data; and (3) concerns about “unknown” actors (hackers and pedophiles). The most common concerns about intimates related to privacy and, in particular, anxiety in relation to using privacy settings. As Josie explained, even though she was “not necessarily talking bad about someone else” mutual friends were a common topic of conversation on social media. Achieving the desired level of disclosure required a detailed understanding of privacy settings, which several female participants were concerned about. There were some differences to privacy concerns based on age and gender. For example, the younger female participants (aged 13–14 years) were more concerned about personal data being accessed and shared by “friends” and “friends of friends,” whereas the older participants (aged 16–17 years) were more concerned about unknown others, such as “hackers” and “identity thieves.”

A concurrent set of concerns related to technology and telecommunications companies tracking participants online and recording their personal data. One participant, Mac, had been using VPNs for “a year or so” to allay his concerns about being tracked online. Concerns around the permanence of social media data were also expressed as another participant, Mary, proclaiming that “what goes online stays online.” This inversion of a well-known quote usually used to maintain discretion among peers (i.e., “what happens in Vegas stays in Vegas”) was used here to convey concerns around the permanence of social media data. As Johanna elaborated, even if a post has been “deleted,” the technology companies had still “got everything” and could use that information at any time in the future.

While participants also expressed concern about other “unknown” actors who they saw as being interested in their social media data, it was hard for participants to articulate who these people actually were. With less concrete examples available, participants described these actors in extreme terms, explaining that they were “pedophiles” and/or “master hackers,” who were, according to one group of boys, motivated by “malicious intent.” Facebook was seen as a platform where one might be more likely to encounter these actors and was a reason *not* to use the platform. Other more predictable concerns around social media use (such as cybersecurity and addiction) were more easily expressed on behalf of others. As Josie explains,

I definitely see how some people would see the downsides of using social media because of things like cyberbullying and stuff like that. But also, some people our age overuse it a bit and so sometimes they get a bit addicted.

Non-Concerns Pre-PDQ. While participants were concerned about maintaining privacy and discretion with intimates, their conversations conveyed a sense of having little control over what friends of friends could do with their social media data. This could be thought of as a pragmatic non-concern; although participants were concerned about what could happen to their data, they realized there was little they could do to maintain control in a networked context. As Johanna explained, “once you send something, it’s no longer yours.” This was accepted as resulting in the fact that the recipient of a post or image “can do whatever they want with it.” Also of little concern to the young people were the actions of so-called “randoms” or unknown people who try to friend or follow them on social media. As Eliza, Josie, and Lucy explain, the technical features of Instagram mean they can easily “block them,” so they posed little threat or concern. Another group explained that as there was no immediate physical threat, such “randoms” were relatively easy to deal with.

Some young people were nonplussed by the collection of their personal data presuming that it held no value to anyone else. As Blair reasoned, “we have nothing of any relevance to anyone . . . we say the most irrelevant, useless stuff.” Another participant, Shane, added that “we pretty much just recommend music,” not perceiving such information as of value to marketers and advertisers. In a similar way, targeted advertising was described as being of little real concern. As Fiona explained, while “it gets rather annoying” to see the same advertisements time and again, these could be easily ignored or overlooked. In fact, many saw targeted advertisements as a rather benign, but necessary, part of using search and social media.

Blindspots and Misunderstandings Pre-PDQ. Several young people were uncertain about what the term “personal data” referred to, with some assuming that these only related to mobile phone use. Similar uncertainties emerged in regard to how social media data are generated and how these are reused and repurposed. For example, Archie assumed that geo-locational data were only generated when he opened the Google maps app on his phone. Other participants were not sure what happened to images and texts once “sent” assuming that these “disappeared.” Indeed, this was a reason why Snapchat, in particular, was so appealing to these participants. However, this feature is not confirmed in Snapchat’s terms and conditions of service.

The workshop interviews also highlighted the young people’s distrust of popular social media platforms such as Snapchat and Facebook. For example, Fiona believed that Snapchat was “run by the government” which was “using its features to track your location at all times.” In a similar way, Chloe, Anna, and Hayley remained suspicious of Facebook and Instagram. They conveyed a belief that the device microphone had to be switched on in order to post content, leading Chloe to think that “it actually listens to your conversations.”

This explanation was used to rationalize her recent experience of being recommended an app by Facebook just moments after speaking about the same app with her mother. Participants had a tendency to exaggerate the capabilities of technology companies, and their interest in participant’s personal lives. This was particularly the case with Facebook, which was generally perceived as untrustworthy. As Joe explained, “it’s a bit murky whether or not you’re on a private account whether it’s actually private.” There were many misunderstandings in regard to Facebook, with several participants assuming that all profiles and posts were public meaning any user can post on any profile page.

Despite these overriding suspicions participants felt toward mainstream social media, many were still using these platforms. Participants spoke of the need to retain a presence on these platforms, as these were important sites for developing social relationships and intimacy with friends. This led to a complex set of reactions in participants in which their sense of “digital savvy” and skepticism was compromised by a sense of social obligation. This was particularly apparent when discussing the terms and conditions of social media platforms. Most admitted to not reading the terms and conditions carefully, with Mac quipping “I just read the button that says, ‘Agree’.” However, this appeared to be said for comic affect, as he later expressed awareness that there were important things included in the terms and conditions. He believed the terms and conditions agreements were deliberately long so that technology companies “can hide stuff in there that you might not like.” Despite this, all participants reasoned that if they wanted to use the service, then they were obliged to agree. As Olivia said, “if you disagree they won’t let you [use the app], so you have to agree.” Rather than a complete misunderstanding, participants were uncertain about the implications of their personal data, and therefore had neither the time nor the inclination to address their concerns.

Young People’s Understandings of Personal Data After the Workshops

New Concerns Post-PDQ. Using the PDQ app and then seeing the data implications of their social media use prompted a number of additional concerns among the young people participating in the research. These new concerns centered primarily around geo-locational tracking and the precision with which it could trace their movements. As Mac said, “It sort of feels like a digital stalker was following me around and it didn’t seem quite right.” Despite consenting to being tracked for the duration of the research project, when presented with the geo-locational data captured through the PDQ app, participants found it “surprising,” “unsettling,” and “creepy.” Most reactions were elicited through the workshops, which demonstrated the capability of PDQ to collate and visualize the meta-data into a list of shops, parks, schools, and center, which participants had visited across the course of the week. Unlike other apps operating on the participant’s mobile

phones, PDQ removed the uncertainty of being tracked. As Mac explained, “Because I know you were tracking me it felt different . . . it put me off using it [PDQ].”

Non-Concerns Post-PDQ. The young people were far less alarmed by the semantic analysis of text and images captured by the app. While some young people were amused by the inaccurate labeling of their emotions in images, others were “annoyed” by the inaccuracy, describing the analysis as “rubbish.” Nevertheless, participants still found it “creepy,” “frustrating,” and “worrying” to be analyzed and typecast in this way. What appeared to be of most concern was that the analysis of text and image was, in Johanna’s words, “opinionated and subjective.” While the misreading of their emotions and sentiments perturbed some young people, others found the inaccuracy reassuring. When confronted with the discursive analysis of text and image, Johanna explained she would prefer to be treated “like a number” rather than “have all these assumptions made about me.”

Using PDQ had little effect on the way the young people felt about targeted advertising. As Matt said, “It doesn’t bother me if I get a random advert for something I’m not interested in.” While Shane simply said, “if you don’t want it, you don’t buy it.” While most young people believed advertisers’ use of social media data for commercial purposes was acceptable, they were still far more concerned about the same information in the hands of others (like “hackers”) who would want to use the data to “make [people] feel worse,” according to Hayley. As she went on to explain, “if [personal data] can be turned into a positive thing then it wouldn’t be so bad.” Indeed, the commercial repurposing of personal data was taken as simply part of contemporary life. As Matt reasoned, “That’s just the way of the world . . . it’s the way things are these days. It’s not a drama. You know about it, but there’s not much you can really do.”

Still Uncertain But More Aware of the Consequences Post-PDQ. While several uncertainties and tensions prevailed, the workshops helped participants develop clearer understandings of the likely reconstitutions of their personal data. In essence, the PDQ app was able to provide an illustration of social media data so that the generation, collation, and repurposing of personal information were made clearer to the young people. This encouraged participants to reflect more carefully on the consequences of what they were doing and instigated discussions about the personal data processes that they would have overlooked in the past. For example, one conversation between two female participants clarified the purpose of cookies, while another participant was inclined to try a false name and birth date when constructing a social media profile. PDQ also provided a more concrete example and explanation for practices that had been suggested to participants in the past. Harry’s parents had told him to set his Facebook to private; however, he was not entirely sure what the purpose of this was:

I was younger then and I asked my parents about it and they said, “Yeah, you can have it as long as you put it on private.” I didn’t understand why back then but I certainly do now after we’ve tested this app out.

Doing Things Differently?

After using PDQ, five of the young people remained unconcerned by the repurposing of their personal data. However, with the realization that their personal details and information were not as private and secure as they first thought, many others expressed emerging concerns. After using PDQ, when asked how they would *like* to feel about their use of social media, many simply said “safe.” At the same time, there was acknowledgment that changing practices in order to be “secure” online was in tension with their typical social media practices, such as presenting a profile and connecting with others. As Blair reasoned, “I want to feel known but I don’t want to feel known by people you don’t really know, like they almost know too much about you just from a few pictures and stuff like that.” For some, being tracked or followed by “unknowns” was described as “scary” and “creepy,” many of the young people were unsure what to do about it. As a result, some resorted to extreme claims, like Eliza, who said, “I’m deleting every app now!”

Using PDQ encouraged several young people to think more carefully about how they might be generating personal data, as John explained, “Now knowing to what extent it can track you in that situation you might be more careful where you use it and how you use it,” while Ricky said, “Having the thought in the back of your head that [third parties] can get more specific things might prompt me to be more careful and cautious when it asks for location or things like that.” At the same time, it was acknowledged that continually monitoring and changing profile and privacy settings required thought and effort, which was too involved for Ty, “I just can’t be bothered to turn off my microphone or my GPS or whatever.” For Ty and two others, any concerns they had about the implications of their personal data were tempered by the idea that few people would actually be interested in their data, or that it would be of any consequence. As Shane said, “it’s not like it’s life or death or whatever.”

Discussion

Unlike some social media research, this study has taken care not to be pejorative about young people’s agency. There is not a normative ideal for personal data practices that we are pushing (see Mathieu, 2016). Put simply, agency encompasses the actions taken by individuals and involves a degree of self-determination. In conducting this research, our goal was to make visible the processes that are often hidden to the end user, so our participants could see how their personal data were being (re)used and make decisions that suited their needs and concerns. Following Couldry (2014), our working

definition of agency also involved creating a longer process of action and opportunities for reflection so that individuals can give “an account of what one has done, even more basically, making sense of the world so as to act within it” (p. 891). As such, this study is interested primarily in what problems users identify, and whether they feel capable of changing their practices as a result of these.

Our research has found that young social media users do have concerns over the ways their personal data are reconstituted and reused in data assemblages. Data assemblages are composed of more than the data system and its infrastructure and include “all of the technological, political, social and economic apparatuses that frames their nature, operation and work” (Kitchin & Lauriault, 2018, p. 8). It is perhaps not surprising then that the ways in which personal data impinge on personal privacy were a particular area of concern for young people. However, the complexity of digital data assemblages, privacy settings, and various “terms and conditions” agreements mean that these concerns are often accompanied by a sense of powerlessness. Of course, it is important to recognize that this is not specifically a “teen” phenomenon—indeed, the issues attendant to personal data similarly overwhelms adults. Yet, our findings suggest that demystifying data processes and making these transparent is a necessary initial step toward raising understandings among young people.

Many of the rationales for not altering individual practices were based on a reluctance to compromise the central tenet of social media—that is, sociality, identity representation, and non-familial community (Haber, 2016). Indeed, the idea of not using social media throws up many concerns and fears for young people. Disclosure of personal data through photographs, posts, and interactions is critical to establishing a digital “presence” on the social network. As Marwick, Fontaine, and Boyd (2017) point out not having a digital presence can influence future job prospects and social opportunities. More significantly, in some instances, non-use of social media has been rhetorically framed as a “political identity statement” (Baumer et al. 2015), meaning there are implications and inferences drawn if one ceases to participate in collectively expected ways.

Our research points to several reasons why data privacy is difficult. First, interacting with others online is a statement of trust and there was a reluctance to jeopardize that through behaving “abnormally.” Rather than disrupt connections and interactions through heightened security settings, most of the young people in our research were inclined to adopt a more permissive attitude toward data privacy. Any concerns they had were offset by the idea that few “people” would be interested in their personal data and that there would only be minor implications through data processing (i.e., targeted advertising). According to van Dijck (2014), this underlying trust in corporate platforms is what has made “datafication”

so widespread. However, unlike van Dijck (2014), our research suggests that this is not an unwitting trust—our participants have a sense that data are reused and repurposed in myriad ways.

Yet the design of social media makes it impossible—if not impractical—to think about data flows. Due to the complex design of social networks, it is very difficult for users to know to whom they are connected through their friends and followers. Depending on a user’s privacy settings, posts can be viewed by potentially thousands of other users, making it “practically impossible . . . to gauge the scope of visibility of the posts throughout the social network” (Jung & Rader, 2016, p. 2). As Gilbert (2012) writes, “Unless explicitly designed into systems, most social information remains hidden inside databases” (n.p.). Without transparency of these connections, it is hard to understand how data flows within the system, let alone manage these through their privacy settings.

It is clear that we cannot simply “trust” companies and businesses to self-regulate. As Peacock (2014) reminds us, the business model of many corporations is based upon information, meaning any form of regulation puts them at an economic disadvantage:

Companies that use big social data in combination with web-tracking technologies store personal data on an unprecedented scale. In an unregulated information market and a *laissez-faire* institutional context, the sheer scale of these operations has introduced new challenges to online user agency. (p. 7)

As such, it is difficult to redress these systemic problems, at least in the short term. A better strategy might be found in lobbying companies to fulfill their corporate responsibility to respect the privacy and security of their customers. More effort should be put into ensuring the terms and conditions are transparent and that renewal of consent is requested, rather than assumed (Custers, 2016).

Technical solutions require competent digital literacies. However, even with competent digital literacies it can be difficult for users to understand and manipulate the technical mechanisms to achieve the appropriate level of disclosure. As Jung and Rader (2016) note,

successfully using the technical mechanisms available in the system to control access to their information is difficult for users, because they often do not understand the correct privacy setting configuration for the level of disclosure they are trying to achieve. (p. 3)

Efforts to cultivate digital literacies also need to encompass the social and ethical aspects of social media, as technical skills alone are not sufficient in preparing young people for the complex situations and decisions they must navigate as part of use.

Conclusion

All of these suggestions point to the underlying need to encourage and support the development of ‘critical data literacies’ among young social media users. As digital technologies continue to permeate everyday life, so too do the opportunities for data extraction. In light of this, it seems prudent to work toward cultivating a new discourse around personal data that impel a more critical disposition toward these issues. Such a discourse might also encourage social media users to think beyond the current strategies and counter-practices to ensure that new problems associated with the rise of Big Data, Internet of Things, and other emerging technologies can be addressed with new practices and strategies.

Despite their youthful intentions and digitally savvy self-perception, these young people were aware that achieving data privacy was difficult. Indeed, Obar (2015) argues that data privacy self-management is a “fallacy” and calls for self-governance as “romantic and impractical” (p. 2). The data assemblage is opaque and complex, putting data privacy and the notion of informed consent beyond the purview of most citizens. As our research has found, managing personal data also requires advanced technical skills and ongoing maintenance. The question then becomes *should* it be up to the individual to ensure their data privacy? Self-responsibilization might be beyond the individual, suggesting that more collective and centralized approaches to data privacy are the only realistic way forward.

Our research has shown that managing and controlling social media data involve social and technical challenges that can be difficult for young people to negotiate. While this might manifest as a form of apathy, it is more likely a sense of powerlessness at the opacity and all-encompassing nature of data assemblages. With limited support and education on how to negotiate these tensions, individuals are left with little choice but to accept the status quo. Clearly, there is a need for further research into how best to support, not only individuals but *all* individuals to become informed and agentic data citizens. While this support will vary according to the demographic in question, demystifying data flows within the data assemblage is an important starting point. An array of critical and creative methods and approaches need to be considered in order to understand and build knowledge of the complex ways in which data can be used to profile and predict individual behavior. Further empirical work is also required to establish which practices are feasible for individuals in managing their data privacy so that academic concerns might align with the realities of everyday data literacies.


Declaration of Conflicting Interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This paper is based on research funded by the auDA Foundation (2016-2017 grants scheme).

ORCID iD

Luci Pangrazio  <https://orcid.org/0000-0002-7346-1313>

References

- Baumer, E., Guha, S., Quan, E., Mimno, D., & Gay, G. (2015). Missing Photos, Suffering Withdrawal, or Finding Freedom? *Social Media & Society*, 1–14. doi:10.1177/2056305115614851.
- Blomberg, J., & Henderson, A. (1990, April). Reflections on participatory design. In Chew, J. and Whiteside, J. (Ed.) *Proceedings of CHI*, 90 (pp. 353–359). Seattle, WA: ACM Press.
- Bødker, S., Grønbaek, K., & Kyng, M. (1993). Cooperative design. In D. Schuler & A. Namioka (Eds.), *Participatory design* (pp. 157–175). Hillsdale, NJ: Lawrence Erlbaum.
- Boyd, D., & Crawford, K. (2012). Critical questions for Big Data. *Information, Communication & Society*, 15, 662–679.
- Couldry, N. (2014). Inaugural: A necessary disenchantment: Myth, agency and injustice in a digital world. *The Sociological Review*, 62, 880–897. doi:10.1111/1467-954X.12158
- Custers, B. (2016). Click here to consent forever: Expiry dates for informed consent. *Big Data & Society*, 3(1), 1–6.
- Dalton C, Taylor L and Thatcher J (2016) Critical data studies. *Big Data & Society*, 3(1), 1–9. doi:10.1177/2053951716648346.
- Denzin, N., & Lincoln, Y. (2005). *The SAGE handbook of qualitative research* (3rd ed.). London, England: SAGE.
- Donovan, G. (2013). *My digital footprint.ORG: Young people and the proprietary ecology of everyday data* (Doctoral dissertation). New York: Faculty of Psychology, City University of New York.
- Ehn, P. (1988). *Work-oriented design of computer artifacts*. Hillsdale, NJ: Lawrence Erlbaum.
- Gilbert, E. (May, 2012). *Designing social translucence over social networks*. Paper presented at the Computer Human Interaction Conference, Austin, TX.
- Gregory, J., & Bowker, G. (2014). The data citizen, the quantified self and personal genomics. In D. Nafus (Ed.), *Quantified* (pp. 211–225). Cambridge, MA: MIT Press.
- Haber, B. (2016). The queer ontology of digital method. *WSQ: Women's Studies Quarterly*, 44, 150–169.
- Helmond, A. (2015). The platformization of the web: Making web data platform ready. *Social Media & Society*, 1, 1–11.
- Jung, Y., & Rader, E. (2016). The imagined audience and privacy concern on Facebook: Differences between producers and consumers. *Social Media & Society*, 2(2), 1–15.
- Kennedy, H., & Moss, G. (2015). Known or knowing publics? Social media data mining and the question of public agency. *Big Data & Society*, 2(2), 1–11.
- Kitchin, R., & Lauriault, T. P. (2018). Towards critical data studies: Charting and unpacking data assemblages and their work. In J. Thatcher, A. Shears, & J. Eckert (Eds.), *Thinking Big Data in geography* (pp. 3–20). Lincoln: University of Nebraska Press.
- Marwick, A., & Boyd, D. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, 16, 1051–1067.

- Marwick, A., Fontaine, C., & Boyd, D. (2017). "Nobody sees it, nobody gets mad": Social media, privacy and personal responsibility among low-SES youth. *Social Media & Society*, 3(2), 1–14.
- Mathieu, D. (2016). Users' encounter with normative discourses on Facebook. *Social Media & Society*, 2(4), 1–11.
- Matzner, T., Masur, P. K., Ochs, C., & von Pape, T. (2016). Do-it-yourself data protection: Empowerment or burden? In S. Gutwirth, R. Leenes, & P. De Hert (Eds.), *Data protection on the move: Current developments in ICT and privacy/data protection* (pp. 277–305). Dordrecht, The Netherlands: Springer.
- Miles, M., & Huberman, A. (1984). *Qualitative data analysis*. Beverly Hills, CA: SAGE.
- Obar, J. (2015). Big Data and the phantom public. *Big Data & Society*, 2(2), 1–16.
- Papathanassopoulos, S. (2015). Privacy 2.0. *Social Media & Society*, 1–2 doi:10.1177/2056305115578141.
- Peacock, S. (2014). How web tracking changes user agency in the age of Big Data: The used user. *Big Data & Society*, 1(2), 1–11.
- Pybus, J., Cote, M., & Blanke, T. (2015). Hacking the social life of Big Data. *Big Data & Society*, 2(2). doi:10.1177/2053951715616649
- Shade, L. R., & Singh, R. (2016). "Honestly, we're not spying on kinds": School surveillance of young people's social media. *Social Media & Society*, 2(4), 1–12.
- Suh, J., & Hargittai, E. (2015). Privacy management on Facebook: Do device type and location of posting matter? *Social Media & Society*, 1(2), 1–11.
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12, 197–208.

Author Biographies

Luci Pangrazio is a postdoctoral research fellow at the Centre for Research for Educational Impact (REDI) at Deakin University, Melbourne. Her research focuses on young people's practices and understandings of personal data. Her "*Young People's Literacies in the Digital Age: Continuities, Conflicts and Contradictions*" (Routledge, 2018).

Neil Selwyn is a professor in the Faculty of Education at Monash University, Melbourne. His current research focuses on issues of digital labor and data-based practices in education. Recent books include "*What is Digital Sociology?*" (Polity 2018) and "*Everyday Schooling in the Digital Age*" (Routledge 2017).