

See discussions, stats, and author profiles for this publication at:  
<https://www.researchgate.net/publication/283895825>

# Deeper data: a response to boyd and Crawford

Article *in* Media Culture & Society · August 2015

DOI: 10.1177/0163443715594105

---

CITATIONS

0

READS

14

1 author:



[André Brock](#)

University of Michigan

13 PUBLICATIONS 103 CITATIONS

SEE PROFILE

# Deeper data: a response to boyd and Crawford

Media, Culture & Society  
2015, Vol. 37(7) 1084–1088  
© The Author(s) 2015  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/0163443715594105  
mcs.sagepub.com  


**Andre Brock**

University of Michigan, USA

## Abstract

Data analysis of any sort is most effective when researchers first take account of the complex ideological processes underlying data's originating impetus, selection bias, and semiotic affordances of the information and communication technologies (ICTs) under examination.

## Keywords

Big Data, critical cultural informatics, critical information studies, data and society, digital sociology, social media and society

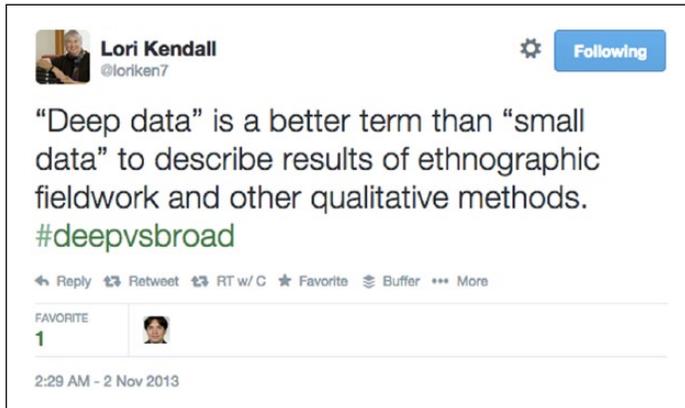


**Figure 1.** #BlackTwitterStudy.

In 2013, Lois Scheidt and I organized a panel for the International Congress of Qualitative Inquiry titled 'Small data in a big data world' as a response to 'Six Provocations for Big Data'. Our panelists presented incredible work conceptualizing new approaches in an age of 'big data' to qualitative social media research, but it was during the Q&A that I came across a term that helped redefine my own critiques of Big

## Corresponding author:

Andre Brock, University of Michigan, 5370 North Quad, 105 S. State Street, Ann Arbor, MI 48109-1285, USA.  
Email: brocka@umich.edu



**Figure 2.** Deep Data.

Data. Lori Kendall, during a spirited Q&A session, offered the term ‘deep data’ as a substitute for ‘small data’ (see also her Tweet below, Figure 2).

Kendall’s contention was that ‘small data’ connotes a lack of complexity, effort, or magnitude; thus positioning qualitative research as being ‘less \_\_\_\_\_ than’ quantitative research. Qualitative researchers’ focus on small groups of participants or singular texts requires significant work – and intensely detailed analysis – to uncover and interpret the many complex meaning-making interactions between entities. To diminish such work by labeling it as ‘small’ reinforces a pejorative positivist view of interpretive inquiry.

Interpretation is at the center of data analysis. Regardless of size or method, all data sets are subject to arbitrarily imposed limitations and bias. While there are methodological tactics available to minimize certain biases, information and communication technologies (ICTs) in and of themselves encourage the imposition of ideological beliefs about rationality, efficiency, and performativity. Data analysis of any sort is most effective when researchers first take account of the complex ideological processes underlying data’s originating impetus, selection bias, and semiotic affordances of the ICTs under examination.

To understand the problems of origination, let us turn to boyd and Crawford’s well-explicated critique of statistical analyses of Twitter data sets. Moving from the shopworn concept of ‘number of tweets’ as an indicator of empirical validity, they question whether social media big data research clearly defines ‘users’, accounts for minimally active users (‘lurkers’), or even authenticates unique users (as opposed to bots or multiple accounts held by a single person or managed accounts) in statistical analyses. This critique isn’t limited to Twitter research, however. I have argued elsewhere (Brock, 2012) that academic conceptualizations of social media (née computer) users tend to draw from functional and instrumental perspectives. ‘Use’ tends to be operationalized as ‘interaction with an interface’; accordingly, ‘social media use’ is understood as ‘interaction with a social media platform or service’. Notice that neither definition encompasses social or

cultural rationales for use. Instead, the technology determines the type of interaction. At best, big data research reframes sociocultural ICT use as ‘patterns’ of use in big data research, while the content or relationships driving that use are black-boxed as a topical concern where  $x$  users respond to  $y$  topic.

This is my contention: to reduce selection bias and unpack semiotic affordances of ICTs, big data research needs to incorporate cultural continuity (Christians, 2007) and reflexivity. ‘Cultural continuity’ refers to an understanding that the ‘the lingual community is understood to be ontologically and axiologically prior to persons’ (p. 440). Lingual here refers to Christian’s argument that communities cohere through language; accordingly, the lingual is not neutral but value laden, and our social bonds are moral claims. From this perspective, big data sets represent cultural, moral, and social choices about technology use as opposed to technologically determined instrumental choices. Nakamura (2013) asked, ‘what is algorithm but ideology in executable form?’ Given that big data research hinges upon the execution of algorithms to gather data and discern patterns, it is entirely appropriate to ask what ideologies are present during the formulation of the inquiry.

A recent example of cultural dis-continuity within big data social media research surfaced when USC Annenberg fumbled<sup>1</sup> its promotion of a preliminary big data analysis of Black Twitter. The researchers planned to capture tweets from *Scandal* viewers, an ABC TV drama produced by a Black woman and featuring a Black female lead. This approach seemed to make an instrumental correlation between Black Twitter users’ information needs and *Scandal* viewers’ social media presence. Black Twitter was not amused (Brown, 2014; Newitz, 2014). One of the more powerful criticisms from these savvy online users argued that Black identity could not be ascertained from ICT practices (Figure 3), in the process indicting this big data approach for its failed functional operationalization of cultural ICT use.



**Figure 3.** DarkSkintDostoyevsky.

I have argued that Black Twitter is but one example of the Black community’s transcendence of the digital divide, signaling material technology adoption, technical literacy mastery, and sophisticated production of user-generated content responsive to the information needs of the community (Brock, 2012). With this in mind, it is not surprising that Black Twitter was able to immediately discern the paucity of a functional perspective on Black ICT use (Figure 4).



**Figure 4.** Zoe Samudzi.

The Tweets above (Figures 1,3 and 4) speak to the ontological aspect of Black Twitter criticism, where concerns about being spoken into existence through the blinkered gaze of academia evoke long-standing fears of exploitation by outsiders. To surface data points like these, I devised an information studies research methodology (critical technocultural discourse analysis or CTDA) for contextualizing social media use within cultural contexts. CTDA frames social media data through triadic technocultural theory and critical race theory, recentring ICT use through the ideologies of their users. It seems facile to assume that other raced and embodied ICT users proffer their cultural identities in exchange for functional mastery of a given ICT. Thus, technoculture and critical race theories work well as analytic frameworks to uncover ideological rationales for technology use.

Returning to cultural continuity, an additional touch point for this response is boyd and Crawford's pointed query as to whether researchers can account for whether their dataset interpretations are appropriate. 'Appropriate' is a loaded term; social scientific concerns about validity or reliability often speak past humanist concerns about cultural context or the application of theoretical frameworks. Moreover, both approaches often espouse normative technocultural beliefs about 'appropriate' technology use, obscuring raced and gendered influences on online activity.

One such example lies within Manovich's (2011) essay on the challenges of big social data. Manovich argues that social media activity is not 'authentic', as online communication is 'often carefully curated and systematically managed' (p. 6). We could, and should, add to this list each platform's mechanical curation, individualized through algorithm, delivering updates and advertisements. The implication that, absent from embodiment, people 'perform' identity through information technologies – as opposed to the authenticity incurred by the increased feedback bandwidth available through face to face communication – rings hollow. As a person of color – and a devotee of Goffman's dramaturgical framework – I can vouch that I am never *not* performing 'self' in every context I enter. To this I add that depending upon context, I am rarely considered authentic by audiences (including but not limited to the academy) even if I correctly enact the signs and commonplaces necessary to fit in. Whether embodied or disembodied, technocultural and cultural ideologies consistently render me as an anomaly.

The preceding is a digression, to be sure, but leads to my final comments. Social media research has been enormously enriched, thanks to access to the large data sets generated through online activity. Because cultural impetuses shape social behaviors, researchers

must avoid making facile claims about online social activity through instrumental analyses. Instead, we should integrate interpretive frameworks drawing from critical cultural theories of technology and identity to evaluate electronic practices and the content generated thereof by socially motivated users. As Christians (2007) writes, 'Because we are all cultural beings – the researched Others, researchers, and the public to whom we communicate – research is not the transmission of specialized data, but in style and content, it reflects moral discernment' (p. 441).

CTDA is only one way to unpack cultural (and moral) rationales for ICT usage. I strongly urge graduate students interested in digital and social media to take courses on the philosophy of technology, as those courses push students to move beyond instrumental philosophies and examine Marxist, phenomenological, and ethical approaches to technology study, design, and use. Finally, cultural affect in ICTs could also be ascertained through institutional mechanisms: by addressing the intellectual and cultural diversity of research teams. While diversity is an oft-abused term, the practical effects of having people from different cultures on a research team studying social phenomena cannot be overstated.

### Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

### Note

1. University of Southern California (USC) neglected to consult with the lead researcher before information about the study, nor did they post her picture on the project website (Chatman, 2014).

### References

- Brock A (2012) From the blackhand side: Twitter as a cultural conversation. *Journal of Broadcasting & Electronic Media* 56(4): 529–549.
- Brown A (2014) Black Twitter being studied by USC, Black Twitter isn't happy, 6 September. Available at: <http://madamenoire.com/466423/black-twitter-studied-usc-black-twitter-isnt-happy/> (accessed 10 October 2014).
- Chatman D (2014) In reply: my reflections on comments about our research on Black Twitter, 4 September. Available at: <http://dchatman3.wordpress.com/2014/09/04/in-reply-my-reflections-on-comments-about-our-research-on-black-twitter/> (accessed 10 October 2014).
- Christians CG (2007) Cultural continuity as an ethical imperative. *Qualitative Inquiry* 13(3): 437–444.
- Kendall L (2013) "Deep data" is a better term than "small data" to describe results of ethnographic fieldwork and other qualitative methods. #deepvsbroad. Available at: <https://twitter.com/loriken7/status/396569822047113216> (accessed 6 August 2015).
- Manovich L (2011) Trending: the promises and the challenges of big social data. In Gold M (ed.) *Debates in the digital humanities*. Minneapolis:University of Minneapolis Press, pp. 460–475.
- Nakamura L (2013) Glitch racism: networks as actors within vernacular Internet theory, 10 December. Available at: <http://culturedigitally.org/2013/12/glitch-racism-networks-as-actors-within-vernacular-internet-theory/> (accessed 18 February 2014).
- Newitz A (2014) What happens when scientists study. *Black Twitter?* 4 September. Available at: <http://io9.com/what-happens-when-scientists-study-black-twitter-1630540515> (accessed 10 October 2014).