The unbearable lightness of information and the impossible gravitas of knowledge: Big Data and the makings of a digital orality

Zizi Papacharissi
University of Illinois at Chicago, USA

Abstract

This essay is written in response and extension to the thoughts offered by danah boyd and Kate Crawford on whether Big Data change how we define knowledge. I suggest that they do not, but they do reinforce and reproduce a form of communicating knowledge that I have been referring to as a digital orality. Online networked platforms, supportive of Big Data and a variety of similar analytical formulations, blend interpersonal and mass storytelling practices variably, offering a reconciliation of primary and secondary orality tendencies and tensions. Literacy, in the form of asking questions about the origins, the textures, and the implications of Big Data, paves the path toward rendering data, small or large, into new modalities of storytelling that a digital orality affords, mastering this orality, and turning these stories into meaningful forms of situated knowledge.

Keywords

Big Data, digital orality, knowledge, literacy, orality, social media

When I first became interested in the relationship between communication and information, conversations about the future of knowledge were rapt with this popular refrain: 'Information wants to be free'. The proclamation had been made by Stewart Brand, to an audience of Internet enthusiasts at the first ever Hacker's Conference in 1984; still it organized how we thought and talked about the information economy well into the next

Corresponding author:
Zizi Papacharissi, University of Illinois at Chicago, 1140 BSB, Chicago, IL 60607, USA.
Email: zizi@uic.edu

decade. The so-called Big Data revolution is the latest stage in a long conversation about how our continuously accelerated ability to distribute and manipulate information is reorganizing our economies, politics, societies, and cultures. The post-industrial society, the information society, and the network society are some of the most prevalent terms for describing the context within which Big Data promise an idealized way to describe, organize, and distribute information about networked realities (e.g. Bell, 1976; Castells, 2000; Wark, 1997). The conversations we have and the refrains we develop in the process absorb and are absorbed in the interrelated hopes and fears about the ability of information to serve as agent for change, empowerment, higher states of knowledge, better democracy, and advanced being. danah boyd and Kate Crawford acknowledge such utopian and dystopian rhetoric triggered by sociotechnical phenomena like Big Data, in their compelling and provocative essay on critical questions presented by the advent of Big Data. I have been tasked with responding to their thoughts on how or whether Big Data change the definition of knowledge and accepted the charge with the hope of contributing to and extending the line of thinking that they develop.

The developing mythology of the capabilities of Big Data attempts to reorganize existing economies of information into economies of knowledge. Do Big Data change how we define knowledge? boyd and Crawford pose this as a question, and in responding to it conclude that what is more important than providing a finite yes or no answer is to continue to investigate the question of how 'harvesters of Big Data might change the meaning of learning, and what new possibilities and new limitations may come with these systems of knowing'. I fundamentally agree, but I find that the response we provide to this question depends largely on how we define the two key terms: Big Data and knowledge. boyd and Crawford offer a meaningful definition of Big Data, which addresses both its analytical premise and promise. They define Big Data as

a cultural, technological, and scholarly phenomenon that rests on the interplay of:

(1) Technology: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets.

(2) Analysis: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.

(3) Mythology: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy. (p. 663)

Understanding Big Data as a phenomenon is central to recognizing both the promise and limitations of Big Data analyses. boyd and Crawford's definition underscores the analytical prowess of Big Data as well as the belief that analytical prowess of this caliber can lead to previously unattainable depths of knowledge. Conventional definitions describe data as sets of values that combine to make up pieces of information, which then combine toward assembling bodies of knowledge. Big Data folklore typically skips the middle step of information and presents Big Data as a mechanism for directly generating

knowledge. The authors present this as a part of the mythological dimension of Big Data, so as to ascertain that, while Big Data is not necessarily capable of de facto knowledge generation, it is certainly and frequently presented as possessing that form of agency.

So do Big Data change how we define knowledge? The literal answer, for me, is no, for the very simple reason that there is no, and there can be no single and universally accepted definition of knowledge. Philosophers have struggled with the question of how we know, what it means to know, and whether it is ever possible to truly know, for thousands of years, dating back to Plato's criteria of belief, truth, and justification, to Descartes claim that 'I think, therefore (I know) that I am', to more advanced and recent distinctions drawn between different theories of knowledge. Because the interplay between objective and subjective ways of knowing renders defining knowledge so complex, and perhaps impossible, the proclamation that Big Data change how we define knowledge seems, at best, odd. And that is why boyd and Crawford want us to think about this question carefully. Thus, they draw from Berry (2011), to note that Big Data offer 'destabilising amounts of knowledge and information that lack the regulating force of philosophy', the epistemology of knowledge, and thus present computationality as its own ontotheology (p. 12).

I would like to extend this line of thinking, by re-inserting the neglected, yet equally abstract, construct of information, in the mix, and by borrowing from the extensive literature on theories of knowledge. Big Data do not change the definition of knowledge. But they do remediate how information is presented, articulated, and visualized, and in doing so, they change the texture of information. And, given that information leads to ways of knowing, then our ways of knowing evolve through new ways of analyzing data, as they always have. But we can get a little more specific than that, and philosophy lends the analytical frameworks for doing so.

Our attempts at defining knowledge inevitably lead us to recognize the inescapable subjectivity of our definitions. Still, they also help delineate ways that we have for arriving at particular formations of knowledge. Examining the ways in which our definitions of knowledge have evolved historically reveal contextual biases of how we understand and interpret the world surrounding us. I suggest that we read the relevance of Big Data within the realm of *situated knowledge(s)*, which emphasize the importance of the subject and the context in perceiving, generating, and reproducing knowledge. Haraway ([1988] 2001) acknowledges the limitations of both human perception and dominant objectivity paradigms and offers this term to describe knowledge as specific to a situation, as presented in the form of a deeply contextual narrative, and as thus offering a more adequate, richer, better account of a world through partiality, and not objectivity. Ultimately, *situated knowledges* help us be part of a better world and more meaningful communities, through 'critical, reflexive relation to our own as well as others' practices of domination and the unequal parts of privilege and oppression that makes up all positions' (p. 172). Big Data present a narrative specific to a particular situation, but they are also part of a greater social reality, within which they are constructed, but one that they also reproduce. They are not devoid of socio-cultural, political, or economic assumptions embedded into the mode of analysis they derive from. They are big, but they are not non-subjective. The sheer size of analysis does not eschew the limitations of subjectivity.

They present iterations of *situated knowledges*, which means that they tell a rich story, but they are also part of a greater story.

Data are signs that combine to present information, that is, symbolic representations of knowledge. It is through these symbolic representations that narratives of knowledge are gradually crafted and attain gravitas through reproduction, recollection, and remediation. Claiming that Big Data change the definition of knowledge implies that somehow this sequence is upset.

That is not the case, at least not yet. Every era is characterized by its own orality or preference for a particular variety of storytelling. It is these traditions of storytelling that eventually amalgamate collective and subjective interpretations of signs and symbols into what we consider relevant and familiar – what comes as close as possible, in a Wittgenstein-ian sense, to what we think of as knowledge. Big Data present a narrative specific to a particular situation, but they are also part of a greater social reality, within which they are constructed, but one that they also reproduce. They are not devoid of socio-cultural, political, or economic assumptions embedded into the mode of analysis that renders them into being. So my proposition is that Big Data make their own contribution to the storytelling practices of our time, and in doing so, they afford narratives of knowledge a unique texture, the texture we might understand as a *digital orality*. They may not change the definition of knowledge, but they do modify how we communicate knowledge.

Signs, be they spoken words, written, symbols, or data, depending on the storytelling conventions of each era, present the means for articulating and communicating knowledge. Each era is driven by its own storytelling conventions, which are frequently considered superior to the storytelling conventions of prior eras. As is commonly known, Socrates despised the written word, for he considered writing to be a deeply inhuman practice and found that true knowledge could only emerge through words as spoken between active human minds engaged in conversation. By contrast, Plato valued Socrates' use of allegory and saw the written word as a way of preserving the power of the spoken word as a pathway to knowledge. In modernity, visual depictions of knowledge come to be considered truer and thus superior to other verbal or written forms of knowledge (e.g. Ong, 1982). Big Data visualizations reflect and amplify these tendencies. Visual depictions of knowledge are not exempt from the same fallacies to which oral and print cultures of communicating knowledge are susceptible, especially when they are often rendered out of audiovisual representations of words. Neither are Big Data analyses, which are in fact meta-analysis of words, put together in a curious and novel blend of interpersonal communication and broadcasting practices – of oral and written storytelling traditions. But they do reinforce and reproduce a form of communicating knowledge that I have been referring to as a *digital orality*.[1]

Orality is used to describe forms of storytelling and knowledge sharing that characterize every epoch. If our goal is to grasp how the technologies of an era afford opportunities for communicating knowledge, then orality describes the form, the texture, the tonality that communication takes on. For example, oral storytelling traditions were driven by a fluidity, spontaneity, and reflexivity in sharing knowledge. Ong (1982) explained that within a primary orality, the nature of the word is not visible but lives in the world of sound. Thus, storytelling 'comes into being in the present even though it

normally may derive variously from a tradition, a past', and evolves as the voice of storytelling changes (Ong, 1995: 1). So knowledge is communicated and constantly evolves as it is communicated. The form of sharing organically changes and evolves as knowledge is reproduced through the spoken word, to listening publics who are aware that this evolving continuity is part of the essence of communicating knowledge. By contrast, the logic and technologies of secondary orality introduced and reproduced the deliberate spontaneity of a writing and print culture, which evolved as electronic and computer-mediated texts accompanied by a secondary orality with a secondary visualism. Ong (1982) had suggested that the logic of a secondary orality is not dependent on *sounding out* stories, in the way that primary orality was. So, in this sense a secondary orality mutes out the subjectivity inherent in voicing this stories, in favor of generating 'technologized', 'permanent', and thus 'silent' stories that distance (Ong, 1982). An example of this form of distance is encountered within the paradigm of journalistic objectivity, which requires narrators to establish objective distance from the story so as to ascertain accuracy and thus electronically reproduce and share verified information. This distance is eliminated in the subjective form of oral storytelling, however, which affectively evolves as it circulates in the oral tradition. Different forms of orality open up different avenues to knowledge, in fact, *to situated knowledges*. No one form of orality leads to truer articulations of knowledge. What is important is to understand the texture of the path to (situated) knowledge each orality opens up.

I suggest that online networked platforms, supportive of Big Data and a variety of similar analytical formulations, blend interpersonal and mass storytelling practices variably, offering a reconciliation of primary and secondary orality tendencies and tensions. Where secondary orality ensured distance, *digital orality* affirms voice, offering a digitally enabled path into the story. Where primary orality emphasized voice, *digital orality* propagates voices while preserving their atomized subjectivity. A *digital orality* is assembled around broadcasting voice, atomized and pluralized. It drives a digital form of storytelling, derivative of the blended conventions of both a primary and secondary orality.

Stories will always be told out of data and have always been pieced together out of data. We learn to tell stories to connect to the world and learn to interpret stories to make sense of the world. Stories vary, depending on the literary and storytelling canon of an era. And communication technologies have always amplified our capacity for storytelling, but in different ways, that invite different forms of affective and cognitive responses. The stories that are told through the technologies and mechanisms of a *digital orality* utilize interconnected and imbricated layers of information, arranged via analytical propositions and algorithmically determined guidelines. As boyd and Crawford argue, 'we must ask difficult questions of Big Data's models of intelligibility before they crystallize into new orthodoxies'. Unless we are familiar with the assumptions embedded in this variety of storytelling – the politics of platforms and the algorithms that afford content a unique texture – our ability to use these stories to make sense of the world surrounding us is compromised. As iterations of *situated knowledges*, Big Data are both thin and thick at the same time – subjective and objective, atomized and pluralized. They promise complete and comprehensive access to information; they cannot deliver complete and comprehensive access to knowledge. They thus oscillate somewhere in between the unbearable lightness of information and the impossible gravitas of knowledge.[2]

Literacy, in the form of asking questions about the origins, the textures, and the implications of Big Data, paves the path toward rendering data, small or large, into new modalities of storytelling that a *digital orality* affords, mastering this orality, and turning these stories into meaningful forms of *situated knowledge*.

Funding

Notes

1. For a discussion of how a digital orality informs and connects networked publics affectively tuning in and out of contemporary issues and affairs, see Papacharissi (2014a). For more on how a digital orality shapes and is shaped by contemporary forms of storytelling, including journalism, as a first draft of history, see Papacharissi (2014b).
2. I refer to the lightness of information as unbearable because it makes information lively and easily circulated, but it also reminds us that information does not yet possess the weight of knowledge. And I think of the gravitas of knowledge as impossible because knowledge must of course possess gravitas while still remaining open to redefinition and evolution.

References

Bell D (1976) The coming of the post-industrial society. *The Educational Forum* 40(4): 574–579.

Berry D (2011) The computational turn: thinking about the digital humanities. *Culture Machine* 12. Available at: http://www.culturemachine.net/index.php/cm/article/view/440/470

Castells M (2000) *The Rise of the Network Society*, 2nd edn. Oxford: Blackwell Publishers.

Haraway D ([1988] 2001) Situated knowledges: the science question in feminism and the privilege of partial perspective. In: Lederman M and Bartsch I (eds) *The Gender and Science Reader*. New York: Routledge, pp. 169–188.

Ong WJ (1982) *Orality and Literacy: The Technologizing of the Word*. London and New York: Routledge.

Ong WJ (1995) Hermeneutic forever: voice, text, digitization, and the 'I'. *Oral Tradition* 10(1): 3–36.

Papacharissi Z (2014a) *Affective Publics: Sentiment, Technology and Politics*. New York and London: Oxford University Press.

Papacharissi Z (2014b) Toward new journalism(s) affective news, hybridity, and liminal spaces. *Journalism Studies* 1–14. Available at: http://www.tandfonline.com/doi/full/10.1080/14616 70X.2014.890328#.VCoYNhbEd9k

Wark M (1997) On the tactic of tactics. *Nettime-l*. Available at: http://www.nettime.org/Lists-Archives/nettime-l-9705/msg00136.html