Ethics Inf Technol (2015) 17:1–9
DOI 10.1007/s10676-014-9357-2

ORIGINAL PAPER

Data barns, ambient intelligence and cloud computing: the tacit epistemology and linguistic representation of Big Data

Lisa Portmess · Sara Tower

Published online: 6 December 2014

Abstract The *explosion* of data grows at a rate of roughly five trillion bits a second, giving rise to greater urgency in conceptualizing the *infosphere* (Floridi 2011) and understanding its implications for knowledge and public policy. Philosophers of technology and information technologists alike who wrestle with ontological and epistemological questions of digital information tend to emphasize, as Floridi does, information as our new ecosystem and human beings as interconnected informational organisms, *inforgs* at home in ambient intelligence. But the linguistic and conceptual representations of Big Data—the massive volume of both structured and unstructured data—and the real world practice of data-mining for patterns and meaningful interpretation of evidence reveal tension and ambiguity in the bold promise of data analytics. This paper explores the tacit epistemology of the rhetoric and representation of Big Data and suggests a richer account of its ambiguities and the paradox of its real world materiality. We argue that Big Data should be recognized as manifesting multiple and conflicting trajectories that reflect human intentionality and particular patterns of power and authority. Such patterns require attentive exploration and moral appraisal if we are to resist simplistic informationist ontologies of Big Data, and the subtle forms of control in the political ecology of Big Data that undermine its promise as transformational knowledge.

L. Portmess (✉)
Department of Philosophy, Gettysburg College, 300 North Washington Street, Gettysburg, PA 17325, USA
e-mail: lportmes@gettysburg.edu

S. Tower
New Lands Farm, Services for New Americans, Ascentria Care Alliance, Gettysburg College '12, West Springfield, MA 01089, USA
e-mail: swtower@gmail.com

Keywords Big Data · Data analytics · Cloud computing · Political ecology · Information theory

Introduction

New technologies such as social media, mobile GPS devices, online banking systems, climate data systems, DNA sequencing, remote sensors, radio-frequency identification (RFID), wireless sensor networks, and ambient data devices generate vast quantities of data that exceed 2.5 quintillion bytes per day (IBM Corporation 2014). An increasing amount of data is machine-generated and unstructured and traces of information that appear random and ephemeral now promise insight through aggregation and analysis in fields such as medicine, agriculture, education and governance. Before 2003, the entire canon of digital information, the *dataverse*, amounted to just 5 exabytes (quintillion bytes) of data. In 2012, nearly 2.5 exabytes were created each day, a number expected to double every 40 months or so (McAfee and Brynjolfsson 2012). *Wired* declared the beginning of the Petabyte Age in 2008, but inevitably petabytes of data have given way to exabytes and data sets of massive historical sweep (Arbesman 2013). Data centers are now being built on a scale unimaginable just a decade ago, to accommodate the enormous growth in data. Range International Information Group is building a 6.3 million sq. ft. super data enter in Langfang, China scheduled for completion in 2016 and expected to be the leader of China's global cloud computing services. Other large data centers have recently opened in Taiwan and Singapore (Google), Sweden (Facebook), Dublin (Microsoft), Bangalore (Tulip Data Services), Wales (Next Generation Data Europe), Utah (NSA) and even larger data centers, and the expansion of existing centers, are planned worldwide for the next decade.

In March 2012 the US government announced a Big Data Research and Development Initiative of $200 million for research aimed at enhancing core technologies of collection, preservation and analysis of data, including grants awarded to the National Institutes of Health for genomic research, US Geological Survey for climate data analysis, and the Department of Defense for autonomous weapons development (The White House 2012). Other projects seek to create to create new tools and methods to extract knowledge from large data sets to accelerate progress in science and engineering research and to build a Big Data workforce. In 2014 the World Economic Forum produced its annual Global Information Technology report entitled "Rewards and Risks of Big Data" urging data-driven innovation and the expansion of the data revolution beyond software engineers to include researchers and policymakers. Such expansion would enable experts to track population health trends, improve understanding of crisis behavior change, map service needs and better predict global changes in supply and demand ahead of major social and political upheavals. The report expresses concern for socioeconomic asymmetry and privacy in a data-driven economy and the increasing challenge of capturing, governing and securing increasingly large amounts of data.

Born of the anomie of the Silicon Valley marketplace, Big Data has thus far largely skirted careful contextualization and moral scrutiny (see however Davis 2012). Yet its linguistic and conceptual representation is richly evocative and suggests emerging quandaries and epistemological issues as the enormity of Big Data and its growth are grasped. Linguistically, the expression *Big Data* frequently seems less descriptive than rhetorical, suggesting new uses and new insights from mining massive data sets yet carrying darker intimations of manipulation and new forms of social control, "a linguistic cousin to the likes of Big Brother, Big Oil and Big Government" (Lohr 2012). This bivalence of Big Data suggests both risk and promise as well as a destabilized rhetoric of information and knowledge. Careful attention to the linguistic and conceptual representations of Big Data reveals emergent ethical concerns in the multiple ways in which the data revolution is experienced, conceptualized and managed. Amidst hype and the promise of corporate bonanzas, Big Data and the new techniques of data-capture, stream computing and text analytics bear heavily on how business decisions are made, consumer behavior predicted, demographic trends interpreted, climate change understood, and objects, processes and people surveilled.

The data revolution has as its ultimate aim objective and dispassionate data analysis that yields usable knowledge. Yet Big Data's promise for knowledge production masks a tacit epistemology that poorly accounts for the pulsing, fragmented, partial and contextual nature of data production. Nor does it account for the intentionality of data harvesting or the dependence of data analytics on complex interpretation rather than on mining and discovery alone. "Information is the ultimate renewable resource. Data reserves not only await harvesting; beneath the surface data are being created, in vast quantities, every day. Finding value from data is much more a process of cultivation than it is one of extraction or refinement" (Thorp 2012). Such data creation cascades from technological innovation and development. Thus data has a time stamp as well as unmistakable positionality; before X-ray technologies were invented no X-ray data could be harvested; before web browsing, no web browsing histories existed. This time-bound nature of data—data that has its place, its distinctive technologies of capture—and data mortality—data that no longer matters, data that has become extinct—are realities lost in the standard epistemology of data. In order to account for ambiguities in the notion of data-mining and data analytics a richer epistemology of Big Data is needed that reveals rather than conceals its social relations and positionality (Harding 2011).

Big Data analytics makes instantaneous access to information a high priority. Yet this focus on anytime everywhere data access conceals the underlying reality of data materiality and the real world servers in which Big Data resides. The tension between the virtual habitat of Big Data and the hidden realities of real world data storage barns shields from view both the political ecology as well as the vulnerability of Big Data to system failure, electricity outages, natural disaster, human error and cybercrime. This vulnerability, as evidenced by the Snowden disclosures, makes plain the positionality and receptivity of Big Data to unforeseen use and exploitation. Moreover, as Big Data increasingly becomes commodified for government, corporate and political analysis, data-mining becomes the province of sophisticated knowers, empowered by institutions with powerful interests who practice what Floridi describes as "the black art called analytics" (Floridi 2010). This casting underscores the darker potentialities and power differentials that emerge when unregulated corporations and unchecked government institutions employ Big Data analytics—whether stream computing analysis, metadata extraction, security analytics, or predictive analytics—without full disclosure or public debate. Nor can Big Data analytics alone bring us closer to an unmediated understanding of the data we harvest, the information it yields and the knowledge we seek from it.

This paper offers an approach to Big Data that explores the tacit epistemology of the linguistic and conceptual representations of Big Data and suggests a richer account of its ambiguities and its multivalence. In *Body-Hacking* *and Data Mining* we examine the ways in which the creative yet unsettled discourse of Big Data conceptualizes the human relationship to Big Data and reflects existential concern through metaphor and semantic paradox. In *Metrics Not Myths* we address the tacit epistemology of Big Data analytics and argue for a conception of Big Data that exists within a matrix of human intentionality and *techné*. Finally, in *Data Is Not The New Oil,* we seek to expose the striking materiality of Big Data masked by both language and corporatized practice, and counter new approaches to the metaphysics of information (Floridi 2011) that neglect the political ecology of Big Data.

***Body-hacking* and *data-mining*: linguistic representations of big data**

Linguistic and visual representations of the *dataverse* have arisen spontaneously as new ideas and abilities demand proper names for reference and context (Hadoop, Apache, Voldemort, data-mining, body-hacking, cloud computing). The metaphorical content of this language is rife with poetic, sardonic and allegorical playfulness that provides a linguistic laboratory of thought for the information revolution. The implicit meanings of these words carry suggestive implications for exploring different ways of envisioning our relationship to emerging information technologies and embody forms of thought and practice significant in efforts to achieve greater control of individuals over their data and protection against fraudulent and exploitive use. The playful linguistic representations of Big Data are revealed in the colorful and childlike industry software packages for processing massive data sets, such as the open source Hadoop, named after the stuffed toy elephant of the developer's son. Similarly, young developers, open-source start-ups, and teams of hacktivists employ boundless creativity in naming their own new enterprises; recent successful software platforms and projects bear whimsical names such as 'squoop,' 'whirr,' 'flume,' and 'giraph,' as well as recycled names such as 'voldemort,' 'kafka,' and 'sensei.' The seeming lightheartedness of these nonce words belies their philosophical import, and their childlike whimsy holds off more serious critique and moral inquiry into the conceptualization and practice of Big Data analytics.

The phrase *Big Data* itself is a trove of suggested meanings for semantic exploration. The nominalization of the word *Big*, a relational predicate that has no fixed meaning, depends entirely on positionality. *Big* is intentionally unsettled and ambiguous; that which constitutes Big Data evolves relative to its medium of use and demand for computational power. Even as our digital media diversify and technological powers multiply, what is Big can never be eclipsed, for nothing can ever become bigger than Big. Moreover, the presence of the polysemic word 'Big' in Big Data has come to evoke a public fear of growing corporate power and declining personal privacy as algorithms learn to predict our next consumer purchase before we do. None of this cultural anxiety is captured by the dictionary definition of 'big'; it is a felt cultural phenomenon concealed by deceptive banality yet couched in dystopic unease.

Big Data is a source of power and of manipulation, which we both seek to possess and struggle against. The existential concern over Big Data often gives representations of Big Data extravagant materiality in vivid metaphors of fire, water and cosmos. The escalation in the volume, variety, and velocity of available data inspires richly evocative and strangely aqueous phrases such as "the exaflood" (Swanson 2007) and "the data deluge" (Anderson 2008). Or the more garish term *data explosion* that social media and the internet of things increasingly deliver. Such vivid terms reflect concern that torrents of Big Data will spectacularly drown us, explosions consume us and "the planetary nervous system" of Big Data absorb us—a graphic visualization of omnishambolic collapse (Smolan and Erwitt 2012). In each of these instances the popular semantics of Big Data suggests a classic and martial dualism. Data analytics thereby is cast as survival knowledge that conquers floods, orders chaos, and stems the information apocalypse.

Yet any hope of an ultimate theory that might provide such survival knowledge is quixotic. "Rather, the creation of data galaxies," David Weinberger (2012) remarks, "has led us to science that sometimes is too rich and complex for reduction into theories. As science has gotten too big to know, we've adopted different ideas about what it means to know at all." In so doing we create models of complexity to map and forecast emergent properties without fully understanding how or why they exist. But experiments abound. For many, Big Data has come to represent the possibility of latent knowledge with far reaching and unknowable potential. The new appearance of hyperdata, data instantaneously captured and used on the spot by producers, consumers and investors, can reveal hyperlocal trends and lucrative opportunities (Hardy 2013). In a high-profile crowd-sourced media project, Rick Smolan and Jennifer Erwitt's *The Human Face of Big Data* (2012) challenge the idea that Big Data is anarchic and devoid of humanity to illustrate the numerous ways in which Big Data will transform society in positive ways, enabling life-changing data-supported technologies to improve our public education, health, safety, and the environment. In a similar vein, the *Quantified Self* movement promises self-betterment through a health conscious 'body-hacking' to track vital bodily rhythms that our conscious brain might otherwise ignore, generating data that can then be used to attain better sleep patterns or a healthier diet, or influence even more subtle changes through "algorithmic nudging" (Morozov 2012). Such projects graft anthropogenic contours onto the competitive corporate future of Big Data, introducing a new interpretive framework that argues for the development of technologies that mediate new social forms that are anything but dehumanizing (Feenburg 1999).

Given the ambiguity of Big Data analytics, it is little surprise to discover how rich and varied colloquial language has become in bending, borrowing, and contorting ordinary words to describe an extraordinary process. Indeed, the colorful variety of descriptions for how we interact with Big Data has striking semantic implications. Common phrases such as *data-capture* transport us to wild jungles of infobeasts where data is ambushed and ensnared, while *data-mining*, *dredging* and *harvesting* suggest something more patiently pursued, static and predictable, though still half-buried and partially unknown. Still other verbs such as *massaging*, *leveraging,* and *curating* suggest a more artful and duplicitous approach that deliberately selects, coaxes, and manipulates specific strains of data to support a particular interpretation or analysis. With different linguistic possibilities available to conceptualize our relationship to Big Data, we experience an incoherent sense of what data represents to us: *what data is for us*. Our metaphorical extension of borrowed words from older worlds erupts in what Ricoeur (1973) has called "emergent meanings," where semantic paradox yields semantic innovation. Yet, the use of metaphor in this way could not be more appropriate, for metaphor is the *techné* of language—it allows language to reach beyond conventional use. Thus we can think and speak intelligibly about data-mining and data-curating, and acquire greater awareness of what it might mean (or obscure) to speak of taming or massaging data.

The lack of a systematic meta-discourse surrounding the polysemy of Big Data reflects the unresolved linguistic laboratory of thought in which the human relationship to Big Data is explored and its character contested. Such open trajectories and indeterminate possibilities nonetheless have a measure of stability, constrained by abiding features of Big Data (Ihde 2012)—its magnitude, its temporality, its dependence on technology and the obscurity of its patterns *for us*. Big Data confronts us with work to be done and new forms of *techné* to be developed. Ultimately, the words we choose in the shifting, living language of information technology are neither trivial nor coincidental. They shape conceptualization and practice—and ultimately the social and legal movements that can give greater control of individuals over their own data, and greater protection against exploitive use.

In the struggle to come to terms with our data-rich social, political and economic environment, we remain poised between polarized views of risk and promise as the language and discourse surrounding Big Data remains unsettled and widely interpretable. In the following section we examine more closely the tacit presuppositions that shape prevailing interpretations of Big Data and argue for a more complex conception of Big Data as human *techné*.

***Metrics, not myths*: the tacit epistemology of big data**

In a recent marketing campaign, Adobe Systems Inc. boldly represents the new promise that Big Data analytics and information sciences bring to the corporate world by declaring all other marketing tactics obsolete. The company posits data analytics as the *only* way to cut through the overly intuitive fumbling of outdated digital marketing strategies and make decisions based on "metrics, not myths" (Adobe Systems Inc. 2012). It is a common refrain among emerging service providers of data analytics and cloud-based management solutions that all businesses must be able to incorporate insights gathered from data collected in real-time in order to stay competitive. Private and hybrid cloud infrastructure now provide diverse options to companies implementing storage and analytic advantages of cloud-based computing (IBM Corporation 2014).

While there is doubtless real economic power and potential to be gained from adopting cloud-based management and data-driven marketing strategies, the smooth language surreptitiously glossing 'data' into 'fact' is troublesome. What is rarely discussed and frequently overlooked by Big Data service providers is the ontological and epistemological multivalence of data: the understanding that data requires human judgment, human interpretation and is itself the result of either direct human technological capture or the *data exhaust*—the infinite stream of phone records, texts, browser histories, GPS data, and other ambient information—of indirect technological capture. Our semantic conceptions of meaningful information preclude the possibility of proto-epistemic data, or pure, unadulterated, unstructured data. On the contrary, data are always already structured, interpreted, and encountered in contexts that inscribe data with a trajectory of influence, one that is inseparable from human intentionalities (Verbeek 2011). While we may seek to hold onto the possibility of cold, objective, analytic facts about the world, data are hot-blooded creatures, collected with a specific intent at a specific time in a specific context, or constitute the unstructured, random results of specific intents. While those intentions may be extensions of other means or ends, they are irreversibly inscribed in data from the start.

The analytical artistry of Nate Silver in the 2008 and 2012 US presidential races inspired admiration for his independent political data analysis, honed on baseball analytics and shaped by his view that careful predictive analysis can distinguish the "signal" (the truth) from the "noise" (what distracts us from the truth) (Silver 2012). Yet key to his successful presidential prediction was his recognition that we perceive "selectively, subjectively, and without much self-regard for the distortions that this causes" (Silver 2012). For Silver, the solution to too much noise and so little signal is to accept greater probability and uncertainty, and acknowledge and correct for the false or distorted assumptions we bring to data analysis. According to Silver, such fallibilism protects against envisioning data as speaking for itself. We gather data and we speak for data. Though Big Data is not the finite data of presidential elections, recognizing the active, constitutive art of data gathering and interpretation is central to understanding Big Data as having multistable trajectories that give rise to not yet explored possibilities (Ihde 2012). For Ihde such multistability affirms the variation within a form of technology and the divergent expression it takes depending on the context in which it develops. Yet it also emphasizes the extent to which technological values can become fully internalized, where new practices are rationalized and the possibility of new and more subtle "forms of control" emerge in the mass consumption of modes of thinking (Marcuse 1991) emerge.

In the 2010 Global Burden of Disease Study, an unparalleled quantity of health data was amassed by the Institute of Health Metrics and Evaluation and its team of 486 authors in 50 countries that compared estimates for disease and disability in 1990 and 2010 among representative populations worldwide. Its method of collection represents a diversity of both disease categories and established metrics for assessing disease and disability (Murray 2012). Diverse challenges in data collection were confronted, including how best to represent the burden of disease based on representative samples, how to quantify disability and how to identify multifactorial diseases. New forms of health data conceptualization, such as *disability-adjusted life years* require inescapable judgment and an art of analysis (Murray 2012). Such Big Data analytics are expensive to undertake, which illustrate why sophisticated techniques of data-mining are carried out largely by governments, research institutions and corporations. The *massification* of data intensifies concern for meaningful pattern detection and avoidance of error. Like other technology-enabled manifestations of the information revolution, technology enabled data-mining of Big Data is a hybridized form of owned knowledge, in which data-mining, like land ownership and management, produces real world value by harvesting widespread crops, and thus cannot be separated from agents and institutions whose decision procedures shape both the data analytics and the knowledge it produces.

Whatever the means of capture or the inadvertent insights gleaned from the random and ephemeral in data exhaust, data at its most fundamental exists as potentiality. Its technologies of capture are human technologies, the patterns we seek have human significance and much of the data of greatest interest for corporations and for governments is human data. In this way the conceptualization of Big Data as exaflood, deluge or explosion is fundamentally wrong. Much as it captures the experience of confronting the extent of Big Data, it obscures its DNA as human *techné*. Big Data, much like the highway systems, manufacturing plants, repair shops, licensing requirements and regulations that have developed from the invention of the automobile, is entangled in a human matrix of intentionality, a socio-technical system of use (Kline 1985). It may confront us as alterity but its bloodlines are our own. Even the data from increasingly intelligent and self-aware devices that sensors capture flows from technologies specific to time and place, and to the human needs and economic interests that manufacture and use them.

The epistemological challenge is thus to reject the belief that Big Data reveals second order truths. Big Data is already too human, its independent facticity an illusion that molds itself to our standing projections of ourselves. We can never fully remove ourselves from our datasets. Like other transformational technologies, Big Data participates in a chiastic structure of co-constitution, whereby "the public is constituted by the technologies that bind it together but in turn it transforms the technologies that constitute it" (Feenberg 2010). It is this co-construction or "entangled hierarchy" (Hofstader 1979) that keeps us locked in feedback loops of increasing abstraction, leaving us to grapple with the possibility that nothing ultimately separates us from the technologies that enframe us. We now look to the political ecology of Big Data in light of this critique, and the ethical urgency in resisting simplistic informationist ontologies of Big Data.

Data is not the new oil: the political ecology of big data

In the process of grappling with the advent of Big Data the metaphorical mapping of this new phenomenon onto familiar concepts and previous defining eras in human history is commonplace. In this case, the aggregation, analysis, and application of an inexhaustible repository of data have frequently been likened to the extraction, refinement, and combustion of crude oil in society. Yet this tendency to portray data as a natural resource akin to fossil fuels—lying in wait beneath prehistoric oceans and geologic formations, waiting there for human exploitation—prevents us from developing a more critical perspective of our production and cultivation of data, a process that is profoundly human. As Thorp (2012) observes about data gleaned from social networks: "personal data is made from the compressed fragments of our personal lives. It is a dense condensate of our human experience." Data are created. Their condensate is grounded in a different form of materiality.

Though we have a standard system and order of magnitude to describe ever increasing volume of data, only the imagination can differentiate easily between terabytes and petabytes, zettabytes and zebibytes. Big Data is a quantity without fixed materiality, though this is not to say that it lacks materiality altogether. Quite the opposite. Much as cloud computing suggests otherwise, Big Data is not stored in the troposphere, but is grounded in overt physicalities, filling massive data barns and server farms in remote cornfields, abandoned mills, frozen tundra and anyplace where electricity is cheap and abundant. The very metaphor of cloud computing resists envisioning data as grounded, and instead represents an imagined opaque and liberate mass, unburdened by the hard mechanics, force of gravity and moral weight which binds conventional tools and technologies. We fail to account for this reality each time we invoke the transcendent amorphous cloud, the ultimate liberation from our wired, mechanical past. The overlooked physicality of Big Data speaks eloquently of our conception of digital and virtual modalities as antipodes of the actual and the real and belies vulnerability to theft, demand overload, battery failure, terrorism and natural disasters. As inforgs, we peruse the internet as naturally and reflexively as walking or breathing. Rarely are we reminded of the complex, capricious and fragile physical infrastructure that keeps the infosphere aloft; the fiber-optic cables, wireless transmitters, servers and routers, power grids, and orbiting satellites, not to mention the industrial centers, mining operations, and manufacturing companies which provide the raw materials for computing hardware, or the socioeconomic and political institutions, trade agreements, copyright laws, and firewalls that mediate our freedom of access. Only when components are compromised do we dimly recognize the illusion of separation between the virtual and the actual and become gripped by the geopolitics of intelligence surveillance and internet balkanization (Brown 2013; Meinrath 2013).

In perhaps the greatest ruse concerning Big Data's immateriality and timelessness, these massive data storage complexes are resource intensive data warehouses, siphoning off precious natural resources to feed an ever growing demand for digital consumption. In the case of other technologies intimately woven into modern life no such ruse exists. The materiality of nuclear power plants, industrial farm animal production, automobile factories, oil platforms and other real world technologies of production, though masked, is well-documented. Yet the nebulous representation of Big Data reveals few hints that Facebook photos, text conversations, email exchanges, financial transactions, climate data and the nearly infinite repository of Google's street view camera shots are stored in data barns as neatly (but not as securely) as socks or silverware in drawers. The tens of thousands of data barns and server farms scattered around the world represent the most striking manifestation of the shocking physicality of Big Data, which most users and consumers experience as placeless, invisible, and free. The internet, and all its associated information and communication technologies, are not typically characterized as requiring space at all—it is not contained by buildings or borders; it floats, transcending boundaries by displaying the same content irrespective of cultural, political, and environmental particulars. Thus, to uncover the grounded, resource intensive, brick and mortar industrial centers that feed our far-away flickering screens is to break through to a richer sense of the materiality of information—and the data sprawl in real world terms of the infrastructure of human information technology.

Consuming many billions of kilowatt-hours per year, data centers constitute a significant burden on power grids and the generating capabilities of the local communities in which they have settled. They consume hundreds of thousands of gallons of water per year to flush through their industrial cooling systems, and require numerous diesel-powered backup generators or banks of lead-acid batteries to protect against power failure. The startling statistics about data center energy and resource consumption challenge widespread illusions about the sleek efficiency and environmental friendliness of the information industry. Most data centers use a fraction of the electricity powering their servers to perform tasks at any given time; the rest is used to keep servers on stand-by in case of a sudden surge. Energy demands have led to increased concern with energy efficiency, green technologies and locations with abundant energy in the construction of new data centers. Google's Finland data center, for example, uses seawater for its cooling and plans to develop wind farms by 2015 in its shift to renewable energy. While some gains in energy efficiency have been made in accordance with rising industry standards, demand for digital services continues to grow exponentially (Mills 2013).

Data centers interact not only with geographical landscapes but with sociopolitical ones as well. In 2011 journalistic investigation put a small town called Quincy in Central Washington on the map, revealing the hardball political games involved in housing and handling six massive data centers (Glantz 2012a, b). The unlikely farming town, with just 6,900 residents, was chosen because of its proximity to the Columbia River and its cheap and abundant hydropower electrical supply. Yet, few were prepared for the legal maneuvers that would become necessary for keeping the powerful industry in check, including regulating emissions and levying fines for errors in power load forecasting. The disproportionate use of electricity by tech giants in Quincy leverages vast inequalities in the system; as the town becomes more economically dependent on the tax revenue generated by these massive data centers, they risk losing their former autonomy and rural identity. Like the tremendous changes wrought to the land and food system by industrial agriculture, or "Big Ag," Big Data generates its own anomalous reconfigurations across social, political, and ecological systems.

Such investigation into the political ecology of Big Data heightens our awareness of the dialectical relationships and power differentials embedded in its materiality. That data are deeply intertwined in social, political and ecological landscapes should not be overlooked, as philosophers of technology have observed of technology and technological artifacts generally (Borgmann 1984; Winner 1986; Ihde 1990; Latour 1999). These relationships play out in many ways, most visibly as rustic cornfields are cleared for urbane server farms, rivers are transformed into apps, and mountains are obliterated for fiber-optic cables so that algorithms can run five milliseconds faster. Yet Floridi (2010), (2011) makes the case that the information revolution is modifying our everyday perspective on the ultimate nature of reality, from a materialist metaphysics in which objects play a decisive role, to an informational one in which objects and processes are de-physicalized and informationalized. "But in advanced information societies, what we still experience as the world offline is bound to become a fully interactive and more responsive environment of wireless, pervasive, distributed a2a (anything to anything) information processes, that works a4a (anywhere for anytime), in real time" (Floridi 2010). The infosphere, he argues, will not be a virtual sphere supported by a genuinely *material* world behind; instead the world itself will be increasingly interpreted and grasped informationally as the infosphere. At this stage reality itself will be understood as information, rather than the infosphere as a way to describe the space of information.

Yet this hopeful and ultimately utopian vision of the information revolution rests on illusion. Even as *inforgs* we increasingly grasp the world *as* information and migrate to information spaces in which our digital relationships unfold and digital commerce flourishes, the information space ultimately is tethered to real world places that deliver the energy that sustains wireless, pervasive, distributed a2a information process. Even the coy *datasexual*—the self-quantified, obsessed and "relentlessly digital" technoself, who masks materiality by possessing perfectly groomed personal data—has an embodied materiality vulnerable to exposure if data management falters (Basulto 2012). Because of the hidden materiality of information and the inescapable materiality of inforgs themselves, the vulnerability of information to real world social and political realities is masked—and concerns over data privacy, regulation and security are muted. We may live in information spaces, with ubiquitous computing and 'smart' objects, but information lives in real world physical structures, engineered to sustain the illusion of the cloud.

Much of the transformation that informationists like Floridi claim for the information revolution derives from what they see as the increasing de-physicalization of mind and the "hypostatization of the conceptual environment designed and inhabited by mind" (2011). In such a conceptual environment, narratives—"values, ideas, fashions, emotion and that intellectually privileged macro-narrative that is the I"—become "information entities" that quietly come to have an ontological status comparable to that of ordinary things (Floridi 2011). This transformational metaphysic, in which ambient intelligence and even the increasing informatization of the human body, including pacemakers, biometric monitors, cancer-fighting nanobots, brain-based WiFi connections" and other technologies of the body, could only emerge in wealthy countries where electricity is abundant, technologies of use widely distributed to give the illusion of universality and permanence, and socio-technical systems invisibly managed amidst powerful national security. That large parts of the world, including impoverished countries where drones strike and sophisticated cyber assaults are unleashed by powerful nations, do not experience invisible socio-technical systems of use suggests that the illusion of living in a de-localized infosphere *reveals* rather than transcends location. For wealthy countries that own the means of data production, and possess the wealth for data analytics that render Big Data useful, it is easy to project technological conditions of use as universal and the information revolution as a decisive e-migration of the species to the infosphere. This projection obscures the vulnerability and materiality of data in technologically sophisticated nations as well as the inverse realities of the data poor and data hungry of the world, for whom nature is not de-physicalized nor the infosphere a comfortable habitation. Everywhere electricity systems are vulnerable to systems failure, everywhere governments can fall and infrastructure collapse, civil war can destroy civilian grids and energy poverty can jeopardize socio-technical systems of use.

Conclusion

No one who has experienced the grounded materiality of information can long sustain the illusion of cloud computing or the luxury of an informationist metaphysic. That wealthy worlds cultivate such an illusion, and support such technological spheres in which—like factories of earlier days—people both consume and produce what sustains corporate and government power is not surprising. What is surprising is the insouciance with which Big Data and the cloudless computing that allow its harvesting are represented and conceptualized, with insufficient regard for the realistic grappling with ethical and public policy issues that the data revolution requires—such as the potential for data misuse by powerful interests, the re-identification of anonymous information and the vulnerability of data to security, privacy and ecological threats.

References

Adobe Systems Inc. (2012). Adobe marketing cloud—Marketing myths debunked. Retrieved from http://www.adobe.com/solutions/digital-marketing.html.

Anderson, C. (2008). The end of theory: The deluge of data makes the scientific method obsolete. *Wired.* Retrieved June 23, 2008, from http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.

Arbesman, S. (2013). Stop hyping big data and start paying attention to 'long data'. *Wired.* Retrieved January 29, 2013, from http://www.wired.com/opinion/2013/01/forget-big-data-think-long-data.

Basulto, D. (2012). Meet the urban datasexual. *Big Think.* Retrieved April 16, 2012, from http://bigthink.com/endless-innovation/meet-the-urban-datasexual.

Borgmann, A. (1984). *Technology and the character of contemporary life: A philosophical inquiry*. Chicago: University of Chicago Press.

Brown, I. (2013). Will NSA revelations lead to the balkanization of the Internet? *The Guardian.* Retrived November 1, 2013, from http://www.theguardian.com/world/2013/nov/01/nsa-revelations-balkanisation-internet.

Davis, K. (2012). *Ethics of Big Data: Balancing risk and innovation*. Sebastopol, CA: O'Reilly Media.

Feenburg, A. (1999). Technology and meaning. In R. Scharff & V. Dusek (Eds.), *Philosophy of technology: The technological condition (2003)*. Hoboken, NJ: Wiley-Blackwell.

Feenberg, A. (2010). Ten paradoxes of technology. *Techné*. Retrieved December 3, 2014, from https://www.sfu.ca/~andrewf/paradoxes.pdf.

Floridi, L. (2010). *Information: A very short introduction*. Oxford: Oxford University Press.

Floridi, L. (2011). *The philosophy of information*. Oxford: Oxford University Press.

Glantz, J. (2012a). Power, pollution and the Internet. *New York Times*. Retrieved September 22, 2012, from http://www.nytimes.com/2012/09/23/technology/data-centers-waste-vast-amounts-of-energy-belying-industry-image.html.

Glantz, J. (2012b). Data barns in a farm town, gobbling power and flexing muscle. *New York Times.* Retrieved September 23, 2012, from http://www.nytimes.com/2012/09/24/technology/data-centers-in-rural-washington-state-gobble-power.html.

Harding, S. (Ed.). (2011). *The postcolonial science and technology studies reader*. Durham, NC: Duke University Press.

Hardy, Q. (2013). Big data's little brother: Start-ups are mining hyperlocal information for global insight. *New York Times*. Retrieved November 10, 2013, from http://www.nytimes.com/2013/11/11/technology/gathering-more-data-faster-to-produce-more-up-to-date-information.html?_r=0.

Hofstader, D. R. (1979). *Gödel, Escher, Bach*. New York: Basic Books.

IBM Corporation. (2014). What is big data? Big data at the speed of business. Retrieved from http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html.

Ihde, D. (1990). *Technology and the lifeworld*. Bloomington, IN: Indiana University Press.

Ihde, D. (2012). *Experimental phenomenologies: Multistabilities*. Albany, NY: SUNY Press.

Kline, S. (1985). What is technology? In R. Scharff & V. Dusek (Eds.), *Philosophy of technology: The technological condition (2003)*. Hoboken, NJ: Wiley-Blackwell.

Latour, B. (1999). *Pandora's hope*. Cambridge, MA: Harvard University Press.

Lohr, S. (2012). How big data became so big. *New York Times*. Retrieved August 11, 2012, from http://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html?_r=0.

Marcuse, H. (1991). *One dimensional man*. Boston, MA: Beacon Press.

McAfee, A. & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*. Retrieved October, 2012, from http://hbr.org/2012/10/big-data-the-management-revolution/ar/1?cm_mmc=SocialHub-_-3271-_-_-5060270310145972803utm_source=socialhub&goback=%2Egde_4575298_member_170973852.

Meinrath, S. (2013). We can't let the internet become balkanized. *Slate*. Retrieved October, 2013, from http://www.slate.com/articles/technology/future_tense/2013/10/internet_balkanization_may_be_a_side_effect_of_the_snowden_surveillance.html.

Mills, M. (2013). Bravo New York times for discovering reality in 'Power, Pollution, and The Internet'. *Forbes.* Retrieved September 25, 2013, from http://www.forbes.com/sites/markpmills/2012/09/25/bravo-new-york-times-for-discovering-reality-in-power-pollution-and-the-internet/.

Morozov, E. (2012). Google should not choose right and wrong. *Financial Times*. Retrieved December 23, 2012, from http://www.ft.com/cms/s/0/ed17b556-49ee-11e2-a7b1-00144feab49a.html#axzz2LhEG0jpl.

Murray, C. (2012). Global burden of diseases, injuries, and risk factors study 2010: Executive summary. *The Lancet.* Retrieved December 13, 2014, from http://www.thelancet.com/themed/global-burden-of-disease.

Ricoeur, P. (1973). Creativity in Language. *Philosophy Today*. Retrieved November 15, 2014, from http://www.scribd.com/doc/24869776/Paul-Ricoeur-Creativity-in-Language.

Silver, N. (2012). *The signal and the noise*. New York: Penguin Press.

Smolan, R., & Erwitt, J. (2012). *The human face of data*. Sausalito, CA: Against all Odds Production.

Swanson, B. (2007). The coming exaflood. *The Wall Street Journal.* Retrieved January 20, 2007, from http://online.wsj.com/article/SB116925820512582318.html.

The White House, Office of Science and Technology Policy. (2012). Obama Administration unveils "Big Data" initiative: Announces $200 million in new R&D investments. Retrieved March 25, 2012, from http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release.pdf.

Thorp, J. (2012). Big data is not the new oil. *Harvard Business Review Blog Network*. Retrieved November 30, 2012, from http://blogs.hbr.org/cs/2012/11/data_humans_and_the_new_oil.html.

Verbeek, P. (2011). *Moralizing technology: Understanding and designing the morality of things*. Chicago: University of Chicago Press.

Weinberger, D. (2012). To know but not understand. *The Atlantic*. Retrieved January 3, 2012, from http://www.theatlantic.com/technology/archive/2012/01/to-know-but-not-understand-david-weinberger-on-science-and-big-data/250820/#.

Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*. Chicago: University of Chicago Press.