Hackathons, data and discourse: Convolutions of the data (logical)

Helen Thornham[1] and Edgar Gómez Cruz[2]

Abstract

This paper draws together empirical findings from our study of hackathons in the UK with literature on big data through three interconnected frameworks: data as discourse, data as datalogical and data as materiality. We suggest not only that hackathons resonate the wider socio-technical and political constructions of (big) data that are currently enacted in policy, education and the corporate sector (to name a few), but also that an investigation of hackathons reveals the extent to which 'data' operates as a powerful discursive tool; how the discourses (and politics) of data mask and reveal a series of tropes pertaining to data; that the politics of data are routinely and simultaneously obscured and claimed with serious implications for expertise and knowledge; and that ultimately, and for the vast majority of hackathons we have attended, the discursive and material constructions of data serve to underpin rather than challenge existing power relations and politics.

Keywords

Hackathons, data, datalogical, discourse, politics, materiality

Introduction

This paper draws together empirical findings from our study of hackathons[1] in the UK with literature on big data through three interconnected frameworks: data as discourse, data as datalogical and data as materiality. We suggest not only that hackathons resonate the wider socio-technical and political constructions of (big) data that are currently enacted in policy, education and the corporate sector (to name a few), but also that an investigation of hackathons reveals the extent to which 'data' operates as a powerful discursive tool; how the discourses (and politics) of data mask and reveal a series of tropes pertaining to data; that the politics of data are routinely and simultaneously obscured and claimed with serious implications for expertise and knowledge; and that ultimately, and for the vast majority of hackathons we have attended, the discursive and material constructions of data serve to underpin rather than challenge existing power relations and politics.

At the same time, (big) data is already, as David Beer has noted: 'an established presence in our everyday cultural lives' (2015: 2) and this means that the material and embodied configurations of data are already normative and quotidian. It also suggests that hackathons can be investigated as a 'datalogical structure' (Clough et al., n.d., 2015) in terms of the events themselves and as particular prototypes pertaining to how innovation and creativity occurs – as increasingly data driven and automated. In relation to hackathon events, these normative and quotidian configurations also highlight a real disjuncture between the way data is discursively constructed, and the way data is variously and unevenly operationalized in the material and embodied processes of the events themselves.

Inhabiting the hack

Hackathons are established methods and 'spaces of innovation' in the corporate sector (see Gómez Cruz and Thornham, 2016) but are also emerging in the

[1]University of Leeds, UK
[2]RMIT University, Australia

Corresponding author:
Helen Thornham, University of Leeds, Leeds LS2 9JT, UK.
Email: h.thornham@leeds.ac.uk

public and third sectors across the globe (see also Irani, 2015; Leckart, 2012; Marlow, 2013; Townsend, 2013 for histories of the hackathons). They have been widely discussed and critiqued within academia and industry[2] and our aim for this paper is to draw on our experiences in order to critically discuss what we perceive as the underlying and connective thread across all the hackathons events we researched: the discursive, technological and material significance of *data.*

For our research, the hackathons we attended were delineated along three axes: (1) Those within the arts and creative sectors, where programmers, coders, hackers, digital artists used hack and 'wreck' methods to explore digital objects as new ways of 'creative' and political disruption (*Octopus, FoAM, Access Space, Digital Labs, MzTek, ODI* and *MadLab*). (2) Those within the public sector, where the model of the hackathon and creative lab has been adopted as means of fast innovation and public engagement (see also Gregg, 2015). An extension of this also includes key user groups (families, young people, disabled groups) who work with coders, programmers and designers to 'solve' a social, health or cultural issue ('Leeds Hack', 'Hack the City', 'Young Peoples Digital Innovation Lab', 'Watch-It', 'Leeds NHS Hack', 'Leeds Data Mill hacks'). (3) Those within the corporate sectors where individuals or groups are invited to develop prototypes or ideas into marketable commodities, usually competitively, through a range of established hackathons methods (Up London, Digital Shoreditch and USTWO – a digital products studio). While these events varied enormously in terms of politics and aims, what was notable for us was not only the reach of the hackathons across many sectors but also the continual feeling of déjà vu we experienced regardless of the event.

We attended just over 20 hackathons in the UK between 2013 and 2016 as participant observers, but we also worked with hackathons organizers and digital artists as they organized and ran hackathons events in order to understand processes behind the events and their position within a wider spectrum of temporality (beyond the discrete event itself). The particular methodological approach we used in our fieldwork was drawn from Pink et al.'s (2015) concept of 'digital ethnography', which is situated in the politics and principles of reflexivity, participation and observation but is also attuned to the (digital) mediatory elements of digital culture (2015: 3). Our aim was to understand the event from the perspectives of those involved, and reflexively consider our own involvement, with a central question around the issue of *innovation:* who claimed it, when and how. Our methodological leanings were towards ethnography and we drew on a wide corpus of research not only in feminist new media theory, ethnography and digital anthropology but also in design,

digital humanities, organizational studies, business management and science and technology studies to anchor our approach (see e.g., Hoholm and Araujo, 2011; Irani, 2015; Suchman et al., 2002). We were also aware of the unique temporal and spatial environment of hackathons and so were also informed by the notion of 'fast' digital ethnography following the work of Hoholm and Araujo (2011), which focuses on 'an emerging object or practice from the inception of an idea to its successful realization (or indeed failure)' (2011: 936). However, we soon came to realize that the object and practice of our research was not innovation, in fact, but *data,* and this posed different methodological and critical questions for us. It led us on two paths: the first was a discussion of the ethnographic material and considerations of methods and practice (written about elsewhere).[3] The second path – encapsulated in this article – was a critical interrogation of the significance of *data* within and beyond these events. Although we reference ethnographic material, then, this paper is a reflective and critical interrogation and not a presentation of the ethnographic material per se.

In what follows, we discuss three interwoven strands that emerged from our research: data as discourse, data as datalogical and data as materiality. These three configurations enable us to elucidate the multiple and oscillatory alignments of data. For example, on the one hand hackathons are only possible *because* of digital technologies and their underpinning datasets, and this is written into the discursive signifiers of the events themselves.[4] This means that data is both a discursive trope with particular affordances and a material 'condition' of the events (Suchman, 2007). At the same time and as we discuss in this paper, data is simultaneously obscured through the infrastructure of the hackathon events where the emphasis is on creativity and innovation. For us, this reconfigures certain power relations and undermines (we argue) the possibility for a deeper interrogation of the politics of data as an operational or conditional force. Given this, hackathons are not just emergent and proliferating events and methods across a range of sectors: they are also a useful lens to interrogate contemporary issues around (big) data not least because they return us to issues around discursive and material automation (see also Bassett, 2016; Clough et al., 2015; Sterne, 2015), metrification and value, the politics of data and a critical inquiry into 'datalogical' turn (see Clough et al., 2015).

Data as discourse

As suggested above, 'data' is a discursive trope with particular affordances for the events themselves and written into the language of the events through an explicit reference to the digital and to particular

cultural moments and politics of digital culture.[5] The second issue is that these constructions of data also obscure a very real discriminatory politics in which non-valuable data ('raw', 'dirty', 'waste' data) is not simply negated, it is actively unsought, dismissed, disappeared. This means there is a politics at the heart of these events that remains unacknowledgeable through the discourse of data. At the same time, because data is constructed as a malleable, benign tool for participants, identifying the discriminatory politics that are mobilized is a difficult task and one that resonates the existing debates within big data about obscured data aggregation, value and waste (see boyd and Crawford, 2012; Clough et al., 2015; Kitchin, 2014b).

These politics can be discussed in a number of ways, but for the purposes of this article, we will focus on two interrelated issues: the first seeks to interrogate how that discursive construct actively glosses over material and lived inequalities such as age, gender, class, ethnicity that are also particular metrics used in the formation of the event (in terms of the demographics who are invited). The second looks at the kinds of data that becomes unavailable – the 'waste' data; data that cannot be 'cleaned' – and so continually fails to frame this data in an operational capacity: not only can 'waste' data *not* frame the subsequent prototypes or products that emerge during the hackathon, such data is also never referenced or acknowledged as potentially framing – they are written out of the process of creation, the discourse that underpins it and the subsequent product that emerges.

For the hack events that were more consciously politically orientated – such as those organized by MadLab or the ODI, the women only hack event, or those run by the NHS that sought intervention into a system of care – participants were carefully invited along key demographic lines. They were chosen because of their particular value: they were women, they represented a certain set of disciplines, they were users of an existing system that needed changing, they had particular expertise or they were the ultimate users of the new system that was sought. To a certain extent then, it is worth noting that the demographic organization of a hack event is also, in a similar fashion to 'clean data', written along a value system where key identity signifiers are aggregated and afforded particular value. In saying this, we are not drawing a direct correlation between clean data and participant demographics. Instead we are simply noting that the processes critiqued in relation to big data (see boyd and Crawford, 2012) and data analytics (Han et al., 2011; Kitchin, 2014b; Miller, 2010) which note the processes of aggregation in the formation of data and along the data points and 'variables that have the most utility' (Kitchin, 2014b: 101) are more widespread practices

than we might at first acknowledge. Similarly just as the power relations behind these processes of aggregation are sometimes obscure, so hack organizers write out their own intervention in their (later) narratives of the socio-technical outcomes of the event and the specific contribution that these chosen demographics contribute *to* the event.

The second point to make here is that many of these demographics could neither find themselves in, nor identify with, the datasets they were invited to use by the organizers. We worked with obese families who could not locate the complexities of their obesity within the datasets they were offered (the datasets pulled together income, class and age for example, but not other health and mental health issues, family relations, debt or education). We worked with not in education, employment or training (NEET) groups who appeared as unemployed and female and in Leeds, but not (also) as being a primary carer or as having a mental or physical disability, for example. In other words, the datasets were always incomplete, yet were presented as 'clean' data to the participants. The issue here is not about the completeness of a dataset; the issue is that certain datasets were valued over others within the parameters of the hackathon (presumably for time, ease, cost, availability, aims of organizers, etc.). At the same time the rationale for valuing certain datasets over others was masked through a discourse of 'cleanliness', which worked to construct those datasets as benign. The participants had to decide whether to actively seek different datasets thus taking up valuable time from the event or to use the available datasets. If they did the latter, the elements that mattered to those demographics (such as being a primary carer or giving a percentage of income support routinely to another family member) remained obscured in whatever they built, or had to be retrospectively and consciously added, while the discourse of 'clean' data remained relatively unchallenged.

Of course, participants have different stakes in the hackathons and indeed the data they use, they have their own politics and aims: the datasets were in no way evenly taken up, but neither were they benign in any way. As many researchers have noted, data is uneven and wrought along political and power relations: data 'depend on hierarchy' (Gitelman and Jackson, 2013: 8); they are 'correlated' in particular ways to construct particular value (Mayer-Schönberger and Cukier, 2013: 70). Yet through the discourse of cleanliness, malleability and openness, the data was simultaneously constructed as the bedrock for and of innovation or creativity (as fuel, a building block, a terrain). We heard this kind of comment at nearly every hackathon we attended: 'This is non-raw data, and it's clean' (Ian Holt, Ordinance Survey.

Digital Shoredich, 2014); 'The data is clean, it's there for you, its ready to use' (Leeds Data Mill Hack, 2013); 'I have a lot of really, really clean data for you. So it's really easy to use' (Bruce Darling, UP London Hackathon, 2013). Seen here, data was constructed as central tools for the events, and better because data is 'ready' to use, speedier and can be automatically and interoperably (computationally) aggregated.

The data was also constructed as open and transparent – and available for anyone to use ('programming is easy, it's just a matter of 'ifs' and where to put them' (Leeds Hack, 2014), or 'all I need is the data, once I have that, its easy' (UP London, 2013). By comparison, 'raw' data required interpretation, sorting, translating: they required *human* agency to make them useable. Our argument here is that the discursive construct of data writes out the possibility of interrogating the silences within the datasets – the discriminations – *and* makes the converse claim that, if anyone can use it, data actually promotes and enforces equality and openness. These discursive constructions of data also reveal huge investment in the creation *of* particular structures in which and through which data can emerge: if 'clean' data is more valuable (or only 'clean' data is *considered* data) this goes precisely against the notion of big data as 'clean' representations of reality (see also Gitelman and Jackson, 2013: 3; Manovich, 2001: 224; Van Dijck, 2014: 4). Discursively, then, there is as much investment in the cleaning of data as there is in trying to demonstrate that data speaks to/for itself. Data as 'clean' demonstrates the results of this investment, whilst also, of course, obscuring the processes behind it (see also Boellstorff, 2013).

During many of the hackathons we attended, participants came up not only against data they could not identify with, then, but they also came up against inaccessible datasets such as land register data, social housing data. The women only hack event (Newcastle, 2015) for example, had an implicit gender politics that had to be forged through the human and material. There was also a concerted effort to 'find' data relating to women in the North East, but the data was unclean, patchy, too 'raw' to use. Instead, the participants ended up using digital archive material or university material that obscured class, age or geographic location, for example. For the subsequent prototypes, this meant that the politics of the data had to be retrospectively and actively imposed through the material and human interventions. In nearly every instance, participants did one of three things: they turned to familiar data and APIs (familiar through their own use – weather data, travel data) or to social media data (twitter or scraping tools that drew on social media data such as Klout[6] or Kred[7]) or they used the APIs made available by organizations at the hackathon (ordinance survey data,

mapping data). For us, this further contributes to a self-fulfilling cycle that perpetually reinforces the notion of data as a priori, self-evident and truthful not least because in these instances, then, it was 'clean' data that shaped what they built, and whatever the original human/machine intention, the data that was not clean did not register either in the process of creativity or the final project.

These issues also reconfigure 'waste' or 'non-representational' data (Clough et al., n.d.; Thrift, 2007) in new ways. Indeed, even if we recognize that in the hackathons, data was generative and conditioned the possibilities for the activities and prototypes (Clough et al., n.d.: 14; Suchman, 2007), we also have to recognize they also created 'silences' (Bowker, 2005: 11–12) as well as the 'absences of relations' (Kitchin, 2014b: 22, see also Vis, 2013) as we discuss below.

Data as datalogical

Many of the hackathon processes described above are built on a wider, long-standing and normative construction of data. Indeed as suggested above, the central underpinning element that enables data to be valued and operationalized in the ways described above relates to the long-term construction of data as 'transparent', 'self-evident', 'the fundamental stuff of truth itself' (Gitelman and Jackson, 2013: 2). In this section, we briefly elucidate some of these issues in order to suggest that this explains some of the power relations and processes described above and their wider iteration beyond data 'itself'. We also suggest that, particularly with the advent of big data, the notion of self-legitimating datalogical structures is gaining traction with a number of implications not only for the subsequent valuing of the human and material but also for wider issues such as expertise and knowledge which are (re)located into the data processes. The irony, of course, is that the conception of datalogical systems as self-legitimating requires a huge amount of complex (and ironically, external) discursive, human *and* material elements to sustain it.

There are a number of accounts (see also boyd and Crawford, 2012; Clough et al., n.d.; Gitelman and Jackson, 2013; Manovich, 2001) that seek to understand the long-term social, political and technological trend towards a valuing of a system over an individual, and the subsequent consequences for contemporary constructions of (big) data as the foundation *of* that system *and* as something that precedes information, knowledge and wisdom (Kitchen, 2014b: 9). David Graeber's (2015) work into bureaucratic systems offers one avenue through which we can begin to understand the long-term shift towards trust and value in both the system per se and systems of metrification

and quantification on which and through which (big) data emerges. Indeed, for Graeber, the history of bureaucracy *is* the story of the Internet (2015: 163) and following this, of big data; his work parallels Mayer-Schönberger and Cukier when the latter talk about how datafication and metrification are age-old conceptual tools of early civilizations (2013: 78). What these theorists have in common is a concern with detailing how human trust and value has shifted from people onto systems, and how this shift has further contributed not only to the contemporary value in/of data and big data in particular, but also how this shift increasingly disables us from interrogating the system, comprehending it, or indeed, intervening in it (see also Berry, 2014). Graeber remarks on how bureaucratic systems have become (like big data) both indispensible *and* incoherent. The difference is that bureaucratic systems have been (relatively successfully) critiqued through (neo)liberalism and right wing politics of the market (Graeber, 2015: 6–7); big data absorbs this critique into its logic as a modus operandi *of* the system.

Following this, the central issue for the purposes of this article is how this long-term valuing and construction of data is operationalized into and conditions subsequent value systems within the hackathon events. As we discussed above, such discursive constructions of data have material consequences in terms of what becomes possible to build, to imagine or to utilize. But this discursive construction of data exacerbates a second critique levelled at hackathons around their politics. Irani has noted, for example, that hackathons are exceptionally good at masking a neoliberal and entrepreneurial politics in a language of innovation and creativity with similar consequences to those noted above (in relation to data) in terms of the socio-technical and material outcomes of hackathon events (2015: 2–3). As she argues, this results not only in the active enforcement of a kind of citizenship as entrepreneurialism, but also that this is imagined and constructed as a positive and productive force for social change (Irani, 2015: 2–3). For the purposes of this article, then such considerations remind us that data operates within and alongside a range of other powerful material and discursive signifiers that similarly 'condition' and shape the event; they also remind us that data is also becoming enmeshed in a wider discourse not only of creativity and innovation but also in the elision of these concepts with that of entrepreneurialism. All of these concepts work to disappear or negate an overt politics under the guise of the hackathon event[8] and its premise of *creativity*.

The second issue, which perhaps draws a more direct parallel with contemporary concerns around data, relates to how the construction of data configures the human or material elements within the hackathon.

Indeed, if we consider the way the human and material are conceived by organizers of events, this becomes clearer. Hackathon events are carefully structured with delineated challenges, discussions, scenarios and research and build times written into them (see also Gómez Cruz and Thornham, 2016; Leckhart, 2012). Irani (2015) describes the design of (particularly corporate) hackathons as wrought with the politics of speed and vision (2015: 19) and this means that, although there may be an overt discussion of technological or material 'play' or even 'failure' (when ideas don't work), in reality and because of the temporal conditions of the events (with strict and rapid times that are conducive to the notion of *fast* innovation); failure, experimentation, labour and play are very minimal activities within the scope of the hack.[9] Instead, as we have observed again and again, discussion is rapidly followed by build – and the build is, because it is both material and human, necessarily the slowest element of an event that prioritizes *speed* not only through the temporal dimensions of the event itself, but also because of the construction of big data as framing the conditions of, and possibilities for, the event. While we explore this in more detail below in terms of the processes of hack events, the issue we want to note here relates to the way disruption, failure, compromise and frustration are all constructed as wholly *human*, as slow, and as a kind of necessary evil of the events. Partly these conceptions emerge from the construction of the events themselves as temporally specific – and as necessarily *fast*. But they also emerge from the conditions introduced by data (as the tool, the bedrock, the grounds on which such events are possible). Data, by comparison to the human and material is conceptualized not only as clean, open, transparent and a priori to inequality and discrimination but data is also *fast*.[10] This is further conflated by what we referred to earlier as self-legitimating datalogical structures that not only position the human and material as separate from such structures, but also as increasingly unnecessary – at least in terms of legitimation or in terms of the generation of meaning or agency. It is *this* issue, of course, that has drawn the most criticism and concern from theorists writing about big data (see boyd and Crawford, 2012; Gitelman and Jackson, 2013; Kitchin, 2014a, 2014b; Manovich, 2001; Mayer-Schönberger and Cukier, 2013).

Indeed, datalogical structures, by comparison with humans, are constructed as dynamically adaptive – they are both in a constantly fluid state *and* adaptively responsive. Datalogical structures generate information and have the capacity to build in response to that information. They are self-referential and self-legitimating (partly because of the necessary obfuscation of the system see Berry, 2014; Graeber, 2015) and even while

the human element responds to them, *they* respond to (through interaction with) the outside (human) environment (see Clough et al., n.d.: 7). This means not only that datalogical structures and systems are self-referential and durable (to use Latour's term 1990) but it also means that the frameworks for measurement and value are forged in the system 'itself'. It does not need legitimation or take-up by human or material agents: and in fact, this process is never as good, as *fast*, as productive *as* the system.

The final point to make here is in relation to the discursive and material process of automation, then, particularly in relation to knowledge and expertise. Indeed, we suggest that the construction of data not only as 'a priori' (Drucker, 2011: 1) to information, bias, value, but also as the underpinning element of self-legitimating datalogical structure, conditions knowledge and expertise in particular ways. If we add wider constructions of smart technology to the conceptions of data already discussed in this article, we not only have 'clean' and 'truthful' data, but we also have clean and truthful data *within* a system that is conceived of as 'anticipatory' (Berry, 2014: 22), 'dynamic' (Van Dijck, 2013), 'networked' (boyd, 2014): as 'smart' (Shepard, 2011). These are all terms that carry ideological weight and seek to elucidate the way that automation is shifting expertise and knowledge away from the human and (back) into the system (Webster, 2014: 244). At the same time, the 'system' is also becoming normatively accepted as illegible (Berry, 2014: 14).[11] Similarly, if expertise, knowledge and legitimacy are increasingly and normatively located within the digital system (rather than with the human), human capacity for intervening into or critiquing it is also shifting – leaving us with a human (slower, less dynamic, less adequate) mode of expertise, that is less capable of, amongst other things, future thinking or planning. In considering, then, the constructions of data *as datalogical*, then, we also have to grapple with a very real politics of data that is not (only) bounded by the experience of the hackathon, but is a much wider phenomenon that collapses our capacity not only of intervention, but of future thinking, imagining and planning; of expertise and knowledge into (to borrow from Clough et al., n.d.) 'the changing parameters of computational arrangements' (Clough et al., n.d.: 15, see also Sterne, 2015).

Data as materiality

The final consideration of this article relates to the concept of data as materiality, which we understand in two key ways. The first is through the transformation of data into a tangible and literal material prototype through embodied socio-technical processes, which

elucidate a range of tensions around what we call mundane and ritualized practices and the discourses of data already discussed above. This offers a more ontological and embodied account of data reconfigured into the specific parameters of a hackathon, as well as raising further critical questions for concepts like creativity, innovation and automation. The second way we understand data as materiality relates to a more traditional concept of technology as 'disciplining' bodies (Foucault, 1991, 1997) following scholars such as Cheney-Lippold (2011) and Nafus and Sherman (2014). Foucault's work explicitly connects metrics and measurement of populations with disciplining and measuring bodies (1997), which correlates if not elides (for the purposes of this article) participant demographics with socio-technical capital (see also Graeber, 2015). In the context of the hackathon with its temporal and spatial frameworks and clear delineation of activities organized into specific events, this notion of disciplining is a clear condition of the events themselves that work to both enable and close down certain embodied interactions and interventions. These conditions, in turn, are centrally framed and conditioned by *data*, in terms of the technical and material conditions for hackathon processes. In relation to this, we suggest that there is a clear correlation between data and certain structures such as innovation and creative structures like those pertaining to (and claimed by) hackathon events, as well as to wider issues such as research, knowledge and expertise. The convolution of the datalogical turn, of course, is that data is increasingly positioned as the condition sine qua non so that despite the socio-technical and material processes discussed in this section, it is *data* that remains *the* powerful trope and condition of (and beyond) the hackathon event.

On the one hand, data is transformed into a tangible and literal prototype through established and familiar practices that are enacted in the hackathon event *as* a process of innovation or creativity. Indeed, within the relatively rigid structures of a hackathon, we found that participants necessarily arrived with established processes, ideas or products that were rolled out in a new environment, rather than entirely negotiated with other participants and arrived at during the event.[12] Indeed, to create a *new* process of prototyping would be beyond the usual temporal and organizational structures of the event. This meant that social dynamics of the hackathon converge with the particular time pressures, resulting in a rapidly emergent dominant idea and process that was conducive to the data available and the socio-technical conditions of the specific event (see Gómez Cruz and Thornham, 2016). Most clearly, this was demonstrated through participants arriving with their own *kit*: wires and systems, software, hardware,

speakers, sensors or recording equipment, weaving looms, projectors, amps. In our observations, this kit is adapted to the available data and the infrastructure of the hackathon to produce for example: knitting code that can be translated into sound code within an established and created software (written by another participant) to be performed at the end of the event; sensors that alter tonal register from male to female (or vice versa), using spoken recordings from available archival data; Apps with established design principles that are made to fit the particular requirements of the event.

There are a number of ways we could consider this. One interpretation relates to how we conceive 'innovation' and 'creativity' as something more in keeping with what Shaun Moores has called 'unreflective, taken-for-granted' (digital) corporeal movement and process (2014: 202) where routine and everyday embodied actions occur in specific places and with specific objects (see also Ingold, 2013; Pink, 2012). Seen here, innovation and creativity are more about our mundane and ritualized practices or relationships with technological objects and processes that are enacted across a variety of contexts (see also Kember and Zylinska, 2012: 120–122), than, for example, novel and negotiated socio-technical methods (Balsamo, 2011). To a certain extent then, we might suggest that 'innovation' is better conceptualized as a *change of context* and that the processes (the infrastructures) of innovation actually remain relatively unchanged. Alternatively, we could argue that innovation occurs in the spaces between the data and human – in the embodied practices of participants – that may also recognize the possibilities of human data agency within the structure of the hackathon. From an organizational perspective, of course, participants arriving with their own kit, expertise or methods further underpins the necessity of ensuring a particular demographic presence at hackathon events not least because there seems a clear correlation between participant demographics and socio-technical capital as measured through the causal relationship of participant demographics and final prototype (see also Graeber, 2015; Nafus and Sherman, 2014). Indeed, organizers of hackathon events (particularly those working with digital artists) reiterated the importance of the inclusion of particular skills and demographics because of how this related to 'innovation' and 'creative' processes'[13] as well as to a wider conception of technology and data as implicitly (through embodied subjective process) *human*. Indeed as one organizer told us:

> our hackathons have all been about treating technology as something that can be dissembled and reimagined not as black boxes that can't be accessed... hackathons for us are about humans reclaiming technology as human, taking it apart, claiming it for yourself. (Uncanny Valley Hack, 2015)

While we are not suggesting these processes or relations *are* transparent or causal, we are suggesting that within a metrified and data-driven landscape, this interpretation has particular traction. It also, in the wider context of big data and self-legitimating structures, suggests that we might be witnessing (as Clough et al. (n.d.) and Graeber (2015) suggest) a much wider appropriation of bureaucratic and datalogical processes into different contexts than we might otherwise note or acknowledge.

A second interpretation of the data as materiality relates to how we conceive processes of 'innovation' and 'creativity' (and even 'imagination' see Gómez Cruz and Thornham, 2016) as primarily data driven (through both the framework of the hackathon and through the methods and kit brought to the event by participants) and what this means for embodied and human subjectivities. Indeed, one argument we could make within the specific temporal and spatial frameworks of the hackathons relates to the ways embodied participation could be understood as a particular form of lived and embodied socio-technical action (even agency) that disrupts the disciplining power of the data through embodied mediation. Indeed, some participants actively used kits as disruptive methods in order to highlight the power relations, assumptions and politics at work within the hackathon. Their participation was in and of itself a disruption to the processes, materialities, discourses and ideologies of the events (see author). In this sense, we could argue that the kits brought to events are disruptive digital–human negotiations that intervene into the process of the hackathon: they are familiar, routine and lived as well as digital, technological and datalogical.

These arguments draw on a range of literature within a longer trajectory of (feminist) STS studies that investigates digital embodiment through concepts like 'intra-action' – entangled in and 'intra' to the digital, technological and discursive (Barad, 2009); digital data–human 'assemblages' (Haraway, 2015; Lupton, 2016: 4), or the 'body multiple' as a site to explore lively data that is 'ingested' or 'emitted' (Mol, 2002, 2008: 402). All of these concepts think through embodied human subjectivity as data in relation to agency and power. Indeed in relation to the hackathon, they offer a powerful framework for enabling embodied data–human agency whereby the data is corporeally and discursively reconfigured through embodied action. But as Deborah Lupton has noted, these actions ultimately generate more data that circulates in a wider 'digital data economy' (2016: 4) beyond the initial corporeal and lived processes. This is not to undermine the

potential of data–human agency but to note its temporal and spatial contextuality. In the case of the hackathon, it seems that that ontological expertise, human everydayness and human–technological processes of design are actually carefully and doubly framed and contextualized by not only the rigid infrastructure of the event itself (and therefore only afforded bounded and embodied agency within the temporal and spatial frameworks of the event), but also by the various conceptions of data discussed in this article (data as discourse, data as datalogical, data as materiality) that continually and ultimately try to (re)position data as *the* powerful force that continue to operate beyond these boundaries of the hackthon in a wider data economy or datalogical system. This isn't always successful – in some cases the participants took their prototypes, created their own narratives of the events, circulated their activities in their own way after the event: they resisted these processes. But even as this occurred, the hackathons and their products were also being made to mean in other environments, where they became evidence of entrepreneurial success, claimed as innovation or innovative practice (see also Irani, 2015; Leckart, 2012, author). Clough et al. understand this as this as (re)organizing embodied practices in relation to the values of the datalogical system, ultimately valuing them (only) in terms of their capability of generating data. Seen here the labour of the participants, their own narratives of their activities and even their prototypes are surplus to, or an aside from, the final product: 'bodily practices themselves instantiate as data, which in turn produces a surplus of bodily practices' (Clough et al., n.d.: 11–12). Given this, and to reiterate our argument above, the convolution of the datalogical turn is that data underpins, frames and generates 'innovation' or 'creative' practices. It is also ultimately positioned as the condition sine qua non – beyond the spatial and temporal lived conditions of a hackathon, *and* as the generator for such bounded activities within a hackathon event that can only ever minimally intervene into the datalogical system which continues to be (drawing on earlier arguments) faster, better, self-legitimating. Our argument for the purposes of this article, then, is that despite the socio-technical and material processes discussed in this section, it is *data* that remains *the* powerful trope and condition of (and beyond) the hackathon event.

Conclusion

This article has explored data, discourse and materiality through the specific lens of the hackathon in order to interrogate the socio-technical and political configurations they engender. What we have demonstrated is that data, as self-fulfilling prophecies and self-

legitimating structures is discursively, operationally and materially constructed as both the basis for, and condition of, creative and innovative practices such as those claimed within the discourses of the hackathon. Perhaps more importantly in the context of this article, these discursive, operational and material constructions of data obscure and mask the enormous effort surrounding data that is necessary to position it *as* self-legitimating and self-fulfilling.

At the same time, and as suggested at the start of this article, data and datalogical structures already have an established presence in our lives. But what this article has also attempted to demonstrate is that the claims made in the relatively new field of big data – around metrification, decision making, around infrastructure and processes of organization – have a long-term trajectory in terms of wider value systems and discursive claims. This means that we not only have to recognize events like hackathons within the discursive, ideological and material parameters of (big) data, but also that hackathons reveal that many of those parameters pre-exist and condition the contemporary discourses of big data as well. This highlights to us that the politics of big data have a long term and normative base beyond datalogical structures, and that the turn towards big data has revealed these politics in new ways that similarly make demands of us (see also Gitelman, 2006; Gitelman and Jackson, 2013).

This last issue raises questions around the kinds of demands that can be made around big data, not least when human interventions are normatively positioned as always already inferior. And our hope is that in recognizing data as discursive, as datalogical and as materiality we can open up a different space where new demands of the digital can be made. But in relation to this, we also have to ask about our own complicity. Indeed, Caroline Bassett has recently critiqued our 'post-digital desire to accept the presence of…technology as a given, but also to put it aside' in the explanation of power (2016: 23). Her comment locates data as already figured within discursive and material power relations (and vice versa) *but* routinely *set aside* in explanations of them. Indeed our interrogation of hackathons reveals a very real discriminatory politics at work in the infrastructure of the event, in the valuing of data, in the participant demographics, in the claims made about the value of such activities. At the same time, these politics are simultaneously negated or disappeared in the datalogical system, or located expressly at the site of the human with lesser interventionist ability. In the end, then, the politics of data that are revealed here as having much wider resonance beyond a datalogical system demand that we consider and question our own roles in the long-term normative constitution of data. This is necessary if we are to move

beyond a simple acknowledgement of sociality as datalogical with its corresponding disempowerment of our own value, and instead seek interventions that move us beyond our own complicity in being primarily and ultimately reconfigured through the values *of* that datalogical system.

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) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: This research was conducted as part of a Digital Economy Community and Cultures Network+ EP/K003585/1.

Notes

1. Although we have participated in a number of variously named and labelled events, and although we are aware of the particular history the term 'hackathon' evokes – sometimes uncomfortably for participants – we will refer to these events collectively as 'hackathons' for the duration of the paper.
2. See also Hunzinger J and Schrock A (eds) (2016) The democratization of hacking and making. *New Media and Society* 18(4). Available at: http://0-nms.sagepub.com.wam.leeds.ac.uk/content/18/4.toc (accessed 18 November 2016).
3. See http://inhabitingthehack.github.io; http://2plqyp1e0nbi44cllfr7pbor.wpengine.netdna-cdn.com/files/2013/01/Staging-the-hackathon_Gomez_Thornham.pdf (accessed 9 September 2016).
4. For example, 'Wreckshops' with the evocation of objects or systems to be unpacked; 'digital labs' with the explicit reference to the digital; 'hackathons' with the reference to hacker culture and durability (see also Coleman, 2012; Irani, 2015; Leckart, 2012; Marlow, 2013 for histories of the hackathons).
5. See http://2plqyp1e0nbi44cllfr7pbor.wpengine.netdna-cdn.com/files/2013/01/Staging-the-hackathon_Gomez_Thornham.pdf (accessed 9 September 2016).
6. https://klout.com/home (accessed 10 November 2016)
7. http://home.kred (accessed 12 October 2016)
8. 'it is just for fun really. There isn't a hidden agenda' (Up London participant 2013); 'this is about sharing, not showing off' (Women only hack participant 2015); 'we're here to create and participate, not compete' (Digital Labs participant 2015), 'The most important thing is that there is a real creative vibe here' (Leeds Data Mill participant 2014); 'they are just really *cute* ideas' (Intel judge, UP London 2013); 'this is about helping the digital artist develop themselves and their own creative practices, it's not an industry-funded event: it's not instrumentalised' (organizer Uncanny Valley event 2015).
9. Or as Leach and Wilson have noted, creativity 'appears in a linear relationship to innovation' (2014: 13).
10. We could also consider here the way that the conception of data as fast also apoliticizes the event not least because the event is always already unable to intervene in the datalogical structures by virtue of being too slow, but also because by comparison to the data on which the hackathon is built, the human and material contribution is always already unnecessary, slow, behind.
11. Which we could also read following Graeber (2015), as a long-term trend towards metrification imposed by bureaucracy.
12. For example:

   we were very careful to invite a range of expertise, not just because this means more interesting projects but also because it means everyone here is an expert. This is not about making the next advanced thing, or about showing off, but about sharing approaches and perspectives and good practice. (Women only hack organizer 11.15)

   we need users *and* designers here to shape a useful product. Their knowledge and experience is crucial. (Leeds Hack Organiser 2014)

13. Or what Clough et al. (n.d.) call 'datalogical system'.

References

Balsamo A (2011) *Designing Culture: The Technological Imagination at Work*. London: Duke University Press.

Barad K (2009) *Meeting the Universe Halfway*. Durham, NC: Duke University Press.

Bassett C (2016) *What Perec Was Looking For: Notes on Automation, the Everyday, and Ethical Writing. In: Wilken R and Clemens J (ed.)* The Afterlives of George Perec. *Edinburgh. Edinburgh University Press*. Unpublished paper 1–27.

Beer D (2015) Productive measures: Culture and measurement in the context of everyday neoliberalism. *Big Data & Society* 2(1): 1–12.

Berry D (2014) Post-digital humanities: Computation and cultural critique in the arts and humanities. *Educause* 49(3): 22–26.

Boellstorff T (2013) Making big data, in theory. *First Monday* 18(10). Available at: http://firstmonday.org/ojs/index.php/fm/article/view/4869/3750 (accessed 15 January 2016).

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

boyd D (2014) *It's Complicated: The Social Lives of Networked Teens*. London: Yale University Press.

boyd D and Crawford K (2012) Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon. *Information, Communication and Society* 15(5): 662–679.

Cheney-Lippold J (2011) A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture and Society* 28: 164–181.

Clough P, Gregory K, Haber B, et al. (2015) The datalogical turn. Unpublished Article. 1–26. Available at: https://www.academia.edu/5986819/The_Datalogical_Turn (accessed 2 December 2015).

Coleman G (2012) *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton, NJ: Princeton University Press.

Drucker J (2011) Humanities approaches to graphical display. *Digital Humanities Quarterly* 5(1): 1–23. Available at: http://beautifuldata.metalab.harvard.edu/wp-content/uploads/2014/08/DHQ_-Digital-Humanities-Quarterly_-Humanities-Approaches-to-Graphical-Display.pdf (accessed 4 January 2016).

Foucault M (1991) *Discipline and Punish: The Birth of the Prison*. London: Penguin Classics.

Foucault M (1997) The birth of biopolitics. In: Rabinow P (ed.) *Michel Foucault, Ethics: Subjectivity and Truth*. New York, NY: The New Press, pp. 73–79.

Gitelman L (2006) *Always Already New: Media, History and the Data of Culture*. Cambridge: MIT Press.

Gitelman L and Jackson V (2013) Introduction. In: Gitelman L (ed.) *Raw Data Is an Oxymoron*. Cambridge: MIT Press, pp. 1–15.

Gómez Cruz E and Thornham H (2016) Staging the hack (athon): Imagining innovation, an ethnographic approach. Working Papers of the Communities and Culture Network+ vol 8. Available at: http://2plqyp1e0nbi44cllfr7pbor.wpengine.netdna-cdn.com/files/2013/01/Staging-the-hackathon_Gomez_Thornham.pdf (accessed 20 November 2016).

Graeber D (2015) *The Utopia of Rules: On Technology, Stupidity and the Secret Joys of Bureaucracy*. London: Melville House.

Gregg M (2015) Hack for good: Speculative labour, app development and the burden of austerity. *The Fibreculture Journal* 25: 183–201. Available at: http://twentyfive.fibreculturejournal.org/fcj-186-hack-for-good-speculative-labour-app-development-and-the-burden-of-austerity/ (accessed 15 January 2015).

Han J, Kamber M and Pei J (2011) *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA: Morgan Kaufmann.

Haraway D (2015) Birth of the kennel: A lecture by Donna Haraway, August 2000. The European Graduate School. Available at: http://www.egs.edu/faculty/donna-haraway/articles/birth-of-the-kennel/ (accessed 26 May 2015).

Hoholm T and Araujo L (2011) Studying innovation processes in real-time: The promises and challenges of ethnography. *Industrial Marketing Management* 40(6): 933–939.

Ingold T (2013) *Making: Anthropology, Archaeology, Art and Architecture*. London: Routledge.

Irani L (2015) Hackathons and the making of entrepreneurial citizenship. *Science, Technology and Human Values* 40: 799–824.

Kember S and Zylinska J (2012) *Life after New Media: Mediation as a Vital Process*. Cambridge, MA: MIT Press.

Kitchin R (2014a) Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): 1–12. Available at: http://bds.sagepub.com/content/1/1/2053951714528481 (accessed 15 January 2016).

Kitchin R (2014b) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.

Leach J and Wilson L (eds) (2014) *Subversion, Conversion, Development: Cross-Cultural Knowledge Exchange and the Politics of Design*. Cambridge: MIT Press.

Leckart S (2012) February 17. The hackathon is on: Pitching and programming the next killer app. *Wired*. Available at: http://www.wired.com/magazine/2012/02/ff_hackathons/ (accessed 4 May 2015).

Lupton D (2016) Digital companion species and eating data: Implications for theorising digital data-human assemblages. *Big Data & Society* 3(1): 1–5.

Manovich L (2001) *The Language of New Media*. Cambridge: MIT Press.

Marlow J (2013) Why do we hackathon? *Wired*. Available at: http://www.wired.com/wiredscience/2013/03/why-do-we-hackathon/ (accessed 5 May 2015).

Mayer-Schönberger V and Cukier K (2013) *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray Publishers.

Miller P (2010) Linked Data and Government. ePSIplatform Topic report no. 7. Available at: http://www.epsiplatform.eu/content/topic-report-no-7-linked-data-and-government-0 (accessed 23 December 2015).

Mol A (2002) *The Body Multiple: Ontology in Medical Practice*. Durham, NC: Duke University Press.

Mol A (2008) I eat an apple. On theorizing subjectivities. *Subjectivity* 22: 28–37.

Moores S (2014) Digital orientations: 'Ways of the Hand' and practical knowing in media uses and other manual activities. *Mobile Media and Communication* 2(2): 196–208.

Nafus D and Sherman J (2014) This one does not go up to 11: The quantified self movement as an alternative big data practice. *International Journal of Communications* 8: 1784–1794.

Pink S (2012) *Situating Everyday Life: Practices and Places*. London: Sage.

Pink S, et al. (2015) *Digital Ethnography: Principles and Practice*. London: Sage.

Shepard M (ed.) (2011) *Sentient City: Ubiquitous Computing, Architecture and the Future of Urban Space*. Cambridge: MIT Press.

Sterne J (2015) *Inaugural Lecture for the Sussex Humanities Lab: Keynote Address*. London: RSA. 3rd December.

Suchman L (2007) Agencies in technology design: Feminist reconfigurations. Unpublished manuscript. Available at: http://goo.gl/tsTTyl (accessed 20 November 2016).

Suchman L, Trigg R and Blomberg J (2002) Working artefacts: Ethnomethods of the prototype. *The British Journal of Sociology* 53(2): 163–179.

Thrift N (2007) *Non-representational Theory: Space, Politics, Affect*. London: Routledge.

Townsend AM (2013) *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. New York, NY: WW Norton & Company.

Van Dijck J (2013) *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford University Press.

Van Dijck J (2014) Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance and Society* 12(2): 197–208.

Vis F (2013) A critical reflection on big data: Considering APIs, researchers and tools as data makers. *First* *Monday* 18(10). Available at: http://firstmonday.org/ojs/index.php/fm/article/view/4878 (accessed 23 December 2015).

Webster F (2014) *Theories of the Information Society*, 4th ed. London: Routledge.