

Data and life on the street

Big Data & Society
 July–December 2014: 1–7
 © The Author(s) 2014
 DOI: 10.1177/2053951714539278
 bds.sagepub.com



Alex S Taylor, Siân Lindley, Tim Regan and David Sweeney

Abstract

What does the abundance of data and proliferation of data-making methods mean for the ordinary person, the person on the street? And, what could they come to mean? In this paper, we present an overview of a year-long project to examine just such questions and complicate, in some ways, what it is to ask them. The project is a collective exercise in which we – a mixture of social scientists, designers and makers – and those living and working on one street in Cambridge (UK), Tenison Road, are working to think through how data might be materialised and come to matter. The project aims to better understand the specificities and contingencies that arise when data is produced and used in place. Mid-way through the project, we use this commentary to give some background to the work and detail one or two of the troubles we have encountered in putting locally relevant data to work. We also touch on a methodological standpoint we are working our way into and through, one that we hope complicates the separations between subject and object in data-making and opens up possibilities for a generative refiguring of the manifold relations.

Keywords

Computation, community, data, design, fieldwork, street life

Take a look at the graphs in Figure 1. They present two fairly typical examples of visualisations charting some kind of data over time. In this case, they are the data a few of us at Microsoft Research in Cambridge (UK) have collected, over 18 hours, of vehicle and bicycle journeys outside our office building.

But why are we collecting this data? To what end? And when and where might it be put to use? To put it another way, how might data like this – the flows of traffic on a mostly residential road – come to matter? Where might it matter and to whom?

It is precisely this line of questioning around data mattering that has motivated a year-long project we (a mixture of social and computer scientists, and designers) are embarking on with the people living and working on Tenison Road. The project, handily called the Tenison Road project, stems from some troubles we have had with all the excitement about data being reported in the popular press and matched by at least some of the claims coming from our academically minded colleagues. It is hard not to open a newspaper, or for that matter a journal issue or conference proceedings, without seeing some quite fantastic stories stemming from a deluge of data of one sort or another.

At one extreme, we have some pretty exaggerated claims envisioning data and especially Big Data as the final answer for understanding just about anything that has been, hitherto, too complex to understand. At the other, we see an abundance of stories predicting that a surveillance society, afforded by the proliferation of personal data and hyperconnectivity, is set to dismantle civic life as we know it. The trouble is that these it must be said not altogether unwarranted concerns (see Solove, 2011) tell us very little about how ordinary folk make sense of data and how they might put it to use if they had a choice. They say very little, if you will, about what data *might be* on the street.

So, the Tenison Road project is in some small and concentrated way a means to examine how data could come to matter in a place and amongst people with some very real ideas and concerns. For a year, eight

Microsoft Research, Cambridge, UK

Corresponding author:

Alex S Taylor, Microsoft Research, 21 Station Road, Cambridge CB1 2FB, UK.

Email: ast@microsoft.com



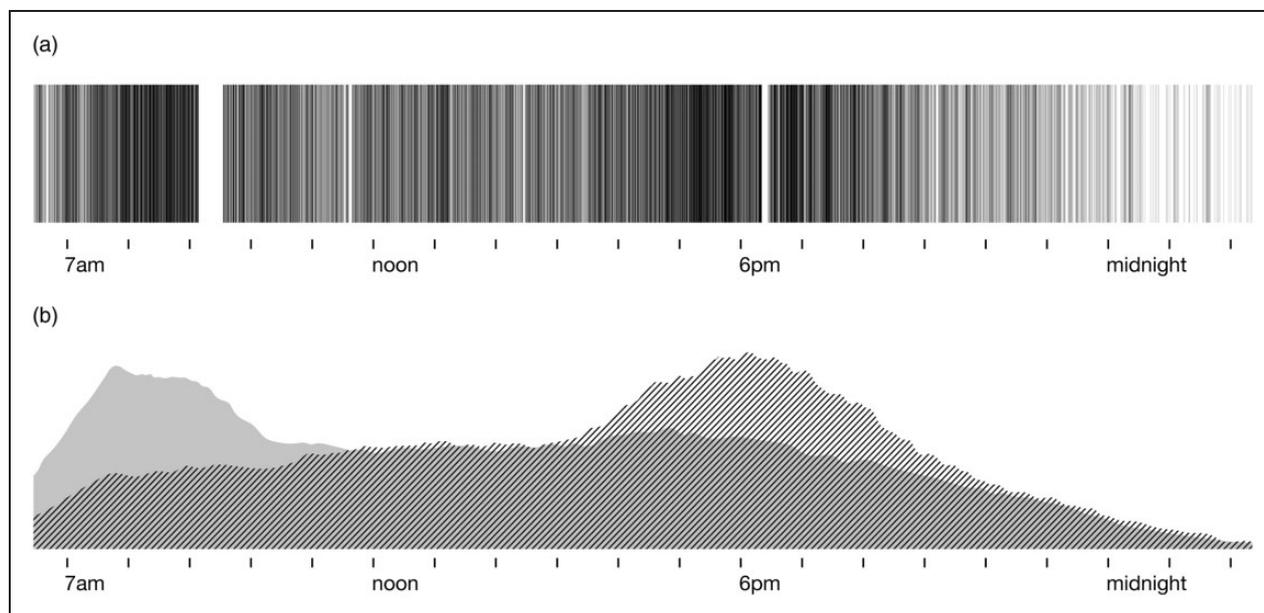


Figure 1. Graphs of (a) density of traffic and (b) direction of traffic along Tenison Road on 25 February 2014. Vertical lines in graph (a) indicate a journey counted so that darker areas are busier than lighter ones (the two relatively wide white gaps represent short periods when the system was down). The uniform grey area in graph (b) indicates number of journeys South, the lighter hashed area indicates journeys North and the darker hashed area is their intersection. Total journeys = 7314 = 3684 (North) + 3630 (South).

of us from Microsoft Research – along with some academic collaborators – are hoping to develop a sense of how data might be materialised and why it might come to matter for those who have a locale in common, people bound not in some seemingly arbitrary way through nodes in a dataset, but by the location they work and live in. As such, the work represents an effort to understand data as coming from somewhere and intended for someone – notably in contrast to how it seems in much of the Big Data rhetoric as ‘raw’ and somehow from *nowhere* and for no one in particular (Mackenzie, 2006). It is then, to use a very particular case, to think about not just technology development but broader themes in social life, such as privacy and data protection; the freedom of access to data; civic participation and engagement; data-related policy making; and generally how data might enable fuller and hopefully better lives for citizens.

So, what is it we are doing exactly? For five months now (beginning in November 2013), we have been meeting, monthly, with some of Tenison Road’s residents – those who have chosen to respond to our flyers, posters and emails (we have over 60 people we are in email contact with and the meetings have between 20 and 30 residents who regularly attend). Over that time, we have collectively spoken about how data might be relevant to Tenison Road and what matters to the ‘community’ (a term that is itself contested). These discussions are giving rise to several threads of activity

including monitoring traffic volumes, noise levels and air quality; building an archive of the road’s social history; and recording particular aspects of contemporary life such as residents’ movements, plant/wildlife, etc.

In gradually building up these disparate forms of data, the aim has been to figure out how they should serve the collective interests of Tenison Road. To this end we have, together with the community, been working with different ways of storing, aggregating and visualising the data. Again, the overarching ambition has been to examine how data might come to matter to the range of people who live and work on a street. Naturally, such an ambition has surfaced tensions around the data we are collecting, how we are collecting it, who has rights to it, and how forms of authority, control, omission, absence, etc. are performed through what we might think of as these collective activities of ‘making data matter’ – or as Wilson (2011: 859) calls it ‘the materiality of code-work’. Furthermore, they have shown the manifold meanings that data can come to have when put to work in everyday life. Let us offer a particular cut through our activities to provide a flavour of the tensions and multiplicities.

The traffic-related data we are beginning to record is, in fact, motivated by significant changes that are underway at the southern end of Tenison Road, where it intersects with the approach to Cambridge’s one and only railway station. At its formal opening in 1889, the road ended some way from the station, separated by

To capture traffic data, we have installed a thermal imaging camera to the roof of the Microsoft Research building at the intersection of Tenison Road and Station Road, in Cambridge (Figure 3(a)). Using heat readings, the camera is able to disambiguate objects from a relatively temperature-uniform background (Figure 3(b)).



Our aim is to develop a range of ways for the data to be analysed and represented, each of which will afford different kinds of interpretations from different types of agents. For example, we are currently building a series of public installations that will display information about speeds, traffic volumes, vehicle types, etc. Figure 3(c) shows one of these visualisation systems in which we are envisioning the data driving remotely controlled, mechanical pie-charts and graphs, creating a spectacle for bystanders on the road to view the 'data' they are themselves producing, as well as ways for them to engage in and comment on it.

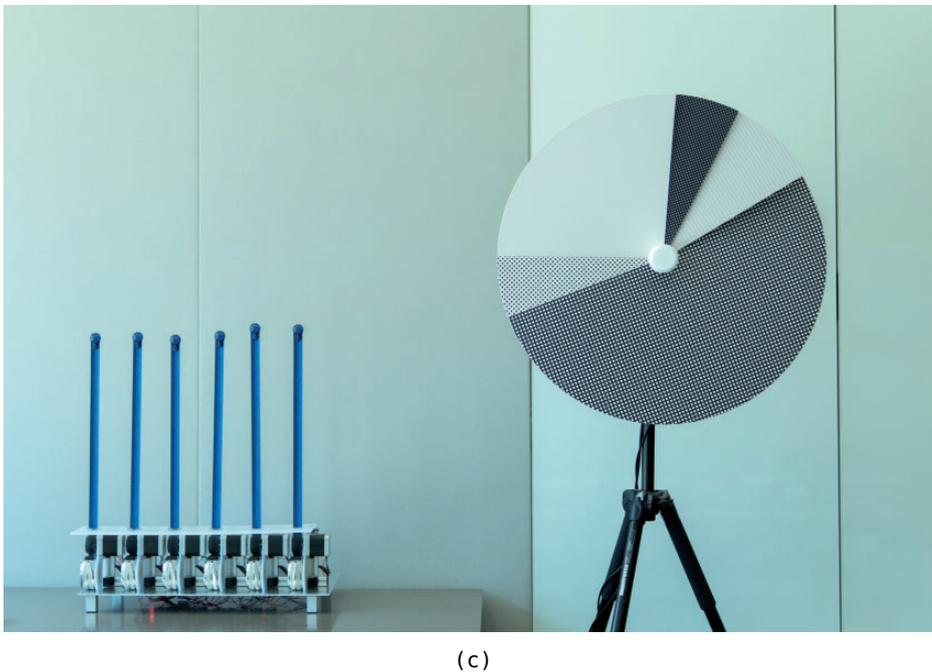


Figure 3. Traffic data collection and visualisation on Tenison Road.

indeed substantial and in aggregate the numbers appear to have been a surprise even to those residents who have been actively involved in our meetings. What has been arguably more enlightening, however, has been how different ideas for data sources have entered into the conversations. The volume and speed of traffic

are, by some, thought not to adequately reflect their worries. In one of our meetings a resident voiced a greater concern for the noise on the street. He went so far as to say he did not especially care how many vehicles used the road or the speed at which they travelled; he just did not want to hear them from inside

his house. Others have voiced a particular worry about air quality. It's evident that even on a single street, the ways in which one data set might be interpreted, and the actions it may seem to imply, vary greatly.

What appears to be at stake in these conversations about data then is the kind of environment the people of Tenison Road believe they live in, and the sort of place they wish to live in. As one of the residents put it in an email correspondence with us:

[W]e have a range of people from those who like to think of themselves as living in a residential back street where traffic should not exist (even if they own two cars) and those who delight in living in a lively cosmopolitan city, albeit one of many commuters according to the traffic data.

What the residents we have worked with have agreed on and appear to form some consensus around is an aspiration for a better quality of life, one that encapsulates a sense of common (or communal) good. Data is – perhaps unsurprisingly given the nature of this project – framed as a means to affect change and improve things such as traffic. As one of the more vociferous residents in the community has phrased it, ‘the purpose of data is action’, and this action is to enable ‘a better quality of life for the community’. Where some of this consensus is complicated, however, is in how exactly the data is understood with respect to collective action and a common good. A better quality of life could mean a quieter street, a safer street, a street where those living across the road are not cut off by a stream of traffic, or a street that is in the midst of a lively cosmopolitan city, with a busy railway station. And each of these necessitates quite different forms of action, collective or otherwise.

To situate the project within these different and, to some extent, competing ideas of a common good, we have initiated a traffic working-group made up of five residents and one Microsoft staff member. As well as making decisions about what exactly to monitor and where, one of the aims of this group is to further involve the community in data collection by, for example, inviting residents to ‘collect’ noise and air quality readings on their premises, data that will, in turn, be fed into a community repository of traffic-related data. The hope is that a network of people may take shape ‘on the ground’ through sharing measuring equipment and a collective involvement in data aggregation and analysis.

The point here is that, at least in the case of our project, data come to be something that matters through the increasing set of relations that are enacted as that data is brought into being and put to use; the mattering is in the ever thickening relations between people and data. The ‘circulation and flow’ of data

that Beer and Borrows (2013) write of and their reference to MacKenzie's (2005) ‘performative circulations’ draw attention to this kind of recursive role data can have with social life. And yet a fluid relationality of this sort raises further troubles for life on the street. We might, for instance, speculate on the possibility of someone on Tenison Road opting out of the traffic monitoring activities – they might see it as a privacy infringement or, as some have, question the efficacy of the exercise. Whatever the case, what would it look like to withdraw or withhold data here, where it is the increasing density of manifold entanglements that give the data significance for the community? The numbers amassed using the thermal camera only begin to have any valence when they are sizeable enough and patterns can be found in their aggregate – the data works because meaning is exercised through the undifferentiated recognition of moving traffic; the patterns, rhythms, changes, etc. are projected in and through the sheer volume of non-descript numbers.

The work involved in the removal of a person's data intervenes in this computational logic by giving primacy to one particular set of figures rather than their aggregates; it introduces the awkward need to identify a unique individual in the data for them then to be removed. To be fair, such an intervention would not ‘break’ the system – despite arguably aiming to achieve the impossible – but it does disrupt how data analytics and computation are understood as efficacious because of their collective and impartial treatment of data. Likewise, we see such disruptions surfacing troubles elsewhere (because, as it has been argued, the computational work is in no way at a different level to the context in which it is applied, but part of the same expansive relational network; Latour et al., 2012). That is, we find there is a disruption to the common good associated with traffic monitoring. With someone opting-out of the data, questions must be asked about what constitutes ‘common’ or ‘collective’ in the case of Tenison Road and how individual exceptions must be accounted for. As with the data and the computation, a logic of coherence through undifferentiated treatment of each individual unravels what it means to assert collective action.

To draw this thread to a close, we want to briefly show that the mattering of data is further complicated when it cuts across different relational assemblages. Somewhat sceptical of our traffic monitoring activities, one resident has voiced questions about the community's role in producing data. She makes a distinction between data ‘coming from us’ and ‘being done to us’. Troubled here are the separations between making data and putting this data to work for the good of the community. In this case, the concern shows how the data-making activities can be understood to be an equally

... I appreciate traffic is a major theme but we all think we don't want to end up being the city council traffic department. . Traffic is about people moving and people moving in the wider aspect is hugely interesting- moving for work, moving for school, moving for fun. we'd all be very up for tracking ourselves and seeing how our movement phenotypes differ.

I think the ideas of archive and record are more interesting and potentially engaging in wider sphere and would be very interested how we could use the tools of today to record life and movement. combining the movement record within the sociological record might be able to make traffic just that little more permanent than the more short term but ever present gripes of urban traffic.

Figure 4. Excerpt from email communication from resident.

important factor in the hoped-for idea of community – that for the data to have some value they must in effect be an intended product of the community.

Also salient is the desire to produce data simply as a way of underpinning an understanding of the street in its current form, located at a point in time, in the context of change. For example, as part of our collective deliberations on traffic, one discussion developed into the possibility of tracking how residents' movements and travel routines might change over the course of the year, reflecting in some way the significant changes afoot in the built environment. To this end, two residents have volunteered to wear location-tracking devices and to feed their geo-temporal data into the larger corpus of traffic data. In itself, this raises further questions about the different qualities of individual versus collective data and how these qualities sit in terms of meaningful aggregation. However, here we want to show what it might mean to read data and consequently perform it in multiple ways.

An email received from a father and his two children (who have regularly attended the community meetings), responding to the traffic monitoring activities and specifically the location tracking (Figure 4), offers an indication of how such data can be associated with different aspects of the street and thus reframed as a record of 'life' on the street. The email asks how the data might be understood in terms of another activity we have begun with Tenison Road, that is, building an archive that represents a social history to the area.

Clearly, this message represents a sophistication in thinking. It refigures the location data in quite a different sense of movement. The data is seen as constitutive of a 'movement record' that locates the to-ings and fro-ings on Tenison Road in wider trajectories of time and social life.

In this refiguring, we can imagine significantly different ways in which the data is materialised. The efficacy of the data is not so clearly judged in terms of collective action and communal good, but rather in how a community hopes to learn about itself and represent itself,

historically. Movement data, but also 'the tools of today', become entangled in particular moments and those moments are located in extended lines of changing human, machine and social relations. Opting in is less a shared conviction towards a common cause as it is marking a juncture of such relations. The data, the tools and the techniques are well and truly put back into life on the street. As such, significant questions are raised about what Beer and Burrows (2013) have characterised as 'by-product data'. They convincingly argue that the data coming out of online transactions and other forms of digital life constitute a by-product that is, in its own right, a highly prized and multifaceted commodity. We don't contest the importance attributed to such by-product data, but seek to show that it is not so hard to find situations in which the modes of collection and aggregation come to matter as much as the data itself. Overemphasising the by-product-ness of the data belies the value its means of production can have in social and (our specific case) street life.

Stepping back, then, the Tenison Road project still feels very much in its early stages (despite being five months in). As those who have conducted community focused projects will attest, it takes a long time to build trusting relationships, glean some sense of the issues at stake, and figure out how to participate in these. At this stage, what we're finding is how productive it has been to quite literally 'go to the street' not only just to discover how things are – how they are materialised and matter for people – but also how they could be. We find, even at the scale of the street, where the data we are dealing with is far from 'big' (in any computational sense), data is folded into social life, and vice versa, so that what comes to matter takes shape in and through the myriad relations. Charts of data such as those traffic graphs above are not in any way outside, behind or below, but knitted into a road's community. If anything, what's remarkable in our collective efforts thus far is how, without question, data has become ordinary, how they have entangled in the efforts of a

community to understand itself as such and decide what else it could be. That is not to say the data does not raise problems in this collective going-on, however. Indeed, it does. By looking to see how the traffic data is materialised in measurements, graphs, partitions, disagreements, etc., we see how questions get raised about its legitimacy and efficacy in both individual and civic participation. Furthermore, such material data practices foreground tensions about how private, public and civic lives are constituted. The materialising of data troubles how these categories are kept separate and, indeed, the reasons of keeping them apart. The very specific ways in which data come to matter, then, are at one and the same time part and parcel of how the relations between a road and its residents – a community – come into being.

We are also coming to see what it means to recognise our entanglements in those self-same troubles that our research efforts are giving shape to. Microsoft being on the street (indeed, part of the looming development), and in some sense part of the community, has been something we have kept in the frame, as it were, to intentionally problematise the supposed rawness of data and the imagined separation between the techniques of data collection and analysis and worldly matters. The traffic records of journeys made on Tenison Road are recorded and analysed by us and re-distributed through our own proprietary technological infrastructures to residents and bystanders so they might take some kind of collective action to improve the quality of life for a community. At the same time, Microsoft stands quite literally as a monolith in this neighbourhood, provoking internal frictions within the community about what exactly constitutes ‘better’. Yet, it is precisely these troubles enacted through this ‘collective going-on’ (Verran, 2001: 159) that we hope to stay with

and productively work through in the coming months. We see these responsive and responsible forms of technoscientific practice (Barad, 2007) as a means to genuinely ask how data come to matter.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

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

References

- Barad KM (2007) *Meeting The Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham, NC: Duke University Press.
- Beer D and Burrows R (2013) Popular culture, digital archives and the new social life of data. *Theory, Culture & Society* 30(4): 47–71.
- Latour B, Jensen P, Venturini T, et al. (2012) ‘The whole is always smaller than its parts’ – A digital test of Gabriel Tarde’s monads. *British Journal of Sociology* 63(4): 590–615.
- Mackenzie A (2005) The performativity of code: Software and cultures of circulation. *Theory, Culture & Society* 22(1): 71–92.
- Mackenzie A (2006) *Cutting Code: Software and Sociality*. New York: Peter Lang.
- Solove D (2011) *Nothing to Hide: The False Tradeoff Between Privacy and Security*. New Haven, CT: Yale University Press.
- Verran H (2001) *Science and African Logic*. London: University of Chicago Press.
- Wilson MW (2011) Data matter(s): Legitimacy, coding, and qualifications-of-life. *Environment and Planning D: Society and Space* 29(5): 857–872.