



This is a repository copy of *Balancing the potential and problems of digital data through action research: methodological reflections*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/82825/>

Version: Accepted Version

Article:

Kennedy, H., Moss, G., Birchall, C. et al. (1 more author) (2014) Balancing the potential and problems of digital data through action research: methodological reflections. *Information Communication and Society*, 18 (2). 172 - 186. ISSN 1369-118X

<https://doi.org/10.1080/1369118X.2014.946434>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Balancing the potential and problems of digital methods through action research: methodological reflections

This paper reports on an action research project with public sector organisations in the UK which experimented with a range of digital methods (social media data mining, social network and issue network mapping, data visualisation), in order to explore their potential usefulness for the public engagement activities of these organisations. We argue that there is a need for small-scale, qualitative studies of cultures of large-scale, quantitative data like ours, to open up spaces in which to reflect critically on the methods with which such data is produced. However, in the paper we highlight the difficulties we had enacting through action research a commitment to both the potential (which might be seen as the action part of action research) and the problems (which might be seen as the research part of action research) of digital methods. Following Hammersley (2002), we suggest that an equal balance between action and research may always be difficult to sustain, in both action research and the use of digital methods. Despite this, we argue that critical discussion of digital methods needs to extend beyond academic spaces – through this move, we suggest, we might open up a space in which to reflect on how these methods might be used for the public good.

Keywords: digital methods, social data, action research, social life of methods, public engagement.

Introduction

Recent excitement about the increased volume, velocity and variety of digital data has led to a flurry of bold assertions about what the analysis of such data might tell us. Claims about the power of data to explain our social world can be found across a broad range of sectors, such that it seems that everyone is ‘doing’ data mining and analysis (Ruppert, Law & Savage, 2013). This data gold rush has been driven by a number of factors, including the growing availability of data on users and their online behavior, as more social activities take place online, and the expansion of social media platforms from which much of this data is taken. These factors are especially relevant to social data, which is the focus of this paper. Social data is produced through people’s online interactions on social media platforms and other web spaces where users communicate with each other (see Manovich, 2011 and Cote, 2014 for more on social data). Such data can be big – that is, characterised by volume, velocity and variety – but it can also be small. It is widely suggested that this data can be used to provide new insights into social networks and relationships, analyse public opinion in real time and on a large scale, and capture people’s actual behaviour as well as their stated attitudes (see for example Thelwall et al 2011).

At the same time, within public sector organisations in the UK, funding is being cut and significant changes are being experienced in the name of austerity (Lowndes & Squires, 2012). In this context, there is a drive to move services online. Digital methods such as social media data mining, sentiment analysis, social network and issue network mapping and visualisation are understandably seen to provide potential solutions to the problem of diminishing resources in organisations like city councils. Public sector organisations are not alone in believing in the potential of digital methods to deliver valuable information about the people they want to engage – academic institutions and commercial companies are equally enthusiastic

participants in the data-driven gold rush. In the social sciences, alongside criticism of the celebratory rhetoric which leads to such beliefs (for example boyd & Crawford, 2012), there is growing acknowledgement of the opportunities for social research that are opened up by the increasingly widespread availability of data of all kinds (Ruppert, 2013; Savage & Burrows, 2009).

But using digital methods is not straightforward. Commentators observe that elite commercial companies like Google, Facebook and Amazon have the best access to data, as well as the best tools and methods to make sense of it (Williamson, 2014). boyd and Crawford (2012) argue that soon only top-tier universities will be able to negotiate access to commercial data sources. Thus digital methods are often mobilised in ways that reproduce existing uneven power relations, and as Bassett asserts (2015 forthcoming), little attention has been paid to how data might be harnessed in the interests of the rest of us, like public bodies, community groups, activists and individuals. In this paper, we focus on one of these groups, public sector organisations, and how they might engage with digital methods. We were motivated to study this topic by a commitment to identifying ways of enabling access to digital methods for those with limited economic means. We were interested in exploring ways of circumventing the threat of a new digital divide based on differential levels of data access (boyd & Crawford, 2012), by examining ways in which resource-poor groups who want to use digital methods for the public good might be able to access them. In the context of current austerity measures, public sector organisations fall into this category: they, like many others, are in danger of ending up on the wrong side of the divide. Furthermore, influenced by the ‘social life of methods’ approach developed by Law and others (Law & Urry, 2004; Law, Ruppert & Savage, 2011), which proposes that methods serve to constitute the things they claim to represent, we wanted to explore the ways in which ‘the public’ might come into being through digital methods. Researching this subject is important, we believe, because the types of publics which emerge through such methods become the basis for decision-making about things like the provision of public services. Thus we wanted to reflect critically on the construction of publics through digital methods.

With these interests in mind, we ran a six-month action research project with three city-based public sector organisations in which we attempted simultaneously to apply digital methods, evaluate their potential use and reflect on their normative consequences. In a prior scoping study, we had identified that our partners were already using some digital methods, but these were not used systematically and whilst partners were keen to do more, resources were limited. Thus they welcomed the opportunity to work collaboratively with us to experiment with and reflect on the digital methods which we describe below. Writing about the impact of digital methods on the social sciences, Ruppert (2013) and Savage and Burrows (2009) suggest that social scientists do just this – that is, maintain an openness to the potential of digital methods, whilst also reflecting on the problems. So we had these aims: to open up a space in which public sector organisations could experiment with digital methods, and to understand their uses, potential value and how they construct public life.

In the paper, we narrate our experience of trying to realise these aims through action research, which was not easy. As our research progressed and the various ways in which data gets shaped became visible, our partners remained committed to using

digital methods, despite their acknowledgement of the ways in which digital data might be partial or non-representative. We describe our role as action researchers in producing this desire below, our critical perspective notwithstanding. Thus our paper highlights the difficulty of enacting through action research a commitment to both the potential and the problems of digital methods, as our feelings of responsibility towards our partners to produce results sometimes eclipsed our attention to critical inquiry. In action research terms, a commitment to exploring the potential of digital methods might be understood as the ‘action’ part of action research, and an interest in evaluating their problems as the ‘research’ part. Following Hammersley (2002), we suggest that sustaining an equal balance between action and research may always be difficult, in both action research and the use of digital methods, as both are motivated by a desire to produce results. In making this argument, the paper demonstrates the importance of small-scale research about large-scale data. Small-scale, qualitative studies open up a space for reflecting on the affordances of digital methods, we argue. This point is made by Baym (2013), when she states that ‘Now, more than ever, we need qualitative sensibilities and methods to help us see what numbers cannot’. We conclude that such sensibilities need to extend beyond academic, social scientific and humanities domains. Digital methods, like action research methods, have their limitations, and whilst this may be easily acknowledged in critical academic spaces, more public discussion of such limitations and how they might be addressed is necessary, in order for digital methods to serve the public good. We hoped to create space for such reflection through our research, but, as we describe below, investing effort in producing results meant that sometimes action was prioritised over research.

In summary, this paper does four things. First, it demonstrates the importance of small-scale, qualitative studies of cultures of large-scale, quantitative data. Second, it illustrates the difficulties of balancing potential/action and problems/inquiry in action research into digital methods. In doing so, it highlights issues with action research methods and issues with digital methods; this is the third contribution of the paper. Finally, we argue that, the difficulties we encountered notwithstanding, critical discussion of digital methods needs to extend beyond the academy, in order to open up a space in which to reflect on how these tools might be used for the public good. Below, we frame our paper in debates about digital methods. After this, we reflect on our methodological decision-making and our action research in practice.

Digital methods

Working with regional public sector organisations to explore how digital methods might be used to enhance their understanding of and engagement with their publics, we were dealing, not surprisingly, with small-scale, local data, not big data as it is widely defined – that is, data that surpasses the ability of standard desktop hardware and software to gather, store and process it. Nonetheless, we draw on some debates about big data here, because they form the context in which our research is situated. We relate these debates to discussion about the social life of methods, which also shaped our research.

Proponents of the ‘social life of methods’ approach, like Law et al (2011), argue that methods are socially shaped. Methods are of the social world. They come into being with purposes and advocates and as such, they have a social life (Law et al, 2011). The sample survey, for example, gained momentum in the UK in the 1960s as

a way of researching social life which could bypass the views of the elite, and as such, was supported – or advocated – by the government of the day. What’s more, Law et al argue, methods have a double social life: they are not only ‘shaped by the social world in which they are located’ but also ‘in turn help to shape that social world’ (2011, p.2). Methods simultaneously represent aspects of the social world and have social effects, constituting the things they claim to represent (Law & Urry, 2004). Because of this, Law et al argue, it is important not simply to think of methods as tools, but to think critically about methods, ‘about what it is that methods are doing, and the status of the data that they’re making’ (2011, p.7). This means treating method as a phenomenon in its own right.

Richard Rogers (2013) uses the phrase digital methods to describe what he calls, to paraphrase him, ‘natively digital methods used for studying natively digital objects’. For Rogers and others at the Digital Methods Initiative (DMI) in Amsterdam where such methods have been developed, this means repurposing the methods of the medium – crawling, scraping, mining, ordering, ranking, mapping, visualising and so on – to analyse digital objects like hyperlinks, tweets, hits, likes, tags and metadata. The purpose of digital methods, according to Rogers, is to answer compelling social and cultural research questions. This stands in sharp contrast to the purpose of commercially developed digital methods tools, which tend to be used as a way for corporations, governments and other elites to generate knowledge about the public, often covertly, without the public’s knowledge. We use the term ‘digital methods’ in a broader sense than Rogers here, to incorporate both the techniques for social and cultural research which he and his collaborators have developed and the more market-oriented tools and approaches used by commercial companies, as they are all available to public sector organisations who may not readily differentiate amongst them. The digital methods that we explored on our project include social network analysis, issue network analysis, exploratory content analysis and visualisation; we say more about these below.

To what extent are digital methods seen as shaping social worlds, following the ‘double social life of methods’ model? The answer to this question is a mixed one. On the one hand, some critics argue that more awareness is needed of the social shaping of big data methods, because there is a ‘widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy,’ argue boyd and Crawford (2012, p.663; see also Helles & Bruhn Jensen, 2013). This quote is representative of a recent flurry of critical commentary on the social life of digital methods. boyd and Crawford’s article, cited above, is one example. and various contributions to a special issue of First Monday on big data (such as Baym, 2013 and Boellstorff, 2013) provide further critique, highlighting the making and shaping of digital data through methods (see also Barnes, 2013 and Gitelman and Jackson, 2013).

In contrast, others suggest that the social shaping of digital methods is already widely acknowledged. Boellstorff (2013) argues that those working with big data understand their datasets as limited representations of the world, conditioned by the theories that frame them. Similarly, Havalais (2013) proposes that ‘No one expects the massive amounts of data collected by the Large Hadron Collider or various shared astronomical instruments to be free from the influence of the process by which they were collected.’ Barnes highlights how science studies has long contested the

assumption that scientific knowledge is the product of a ‘disembodied universal logic’ (2013, p.298), and suggests that such a position clearly applies to big data. These writers suggest that social scientists already understand the double social life of digital methods; whether these understandings extend beyond the academy was something that we explored in our research.

But some data sociologists argue that acknowledging the double social life of digital methods – the ways in which they simultaneously are socially shaped and shape the social – is only part of the story. Social scientists also need to take seriously the possibilities which digital methods open up. Ruppert (2013) proposes that social scientists need to connect with digital methods and data practitioners, through ‘cross-sectoral engagements between the social sciences, industry, government and business’ (see also Ruppert et al, 2013), something which we attempted to do, through action research, on our project.

Digital data tools and techniques are frequently developed in corporate contexts and accompanied by commercial rhetoric which aims to convince a wide range of actors that they need digital data. The limited access to tools that results from their development in restricted commercial environments means that corporations are dominating the field of digital data analytics, and this results in the reproduction of unequal power relations referenced in the introduction. As Savage and Burrows (2007, 2009) suggest, there is, therefore, a need to campaign for access to data for the rest of us. But even when access is possible, the aims of digital methods are distinct across sectors, and commercially-developed tools may not be appropriate for all (Baym, 2013). Yet despite concerns about the appropriateness of digital methods tools for all actors, there is widespread interest amongst non-commercial groups in exploring the possibilities that digital methods and digital data might enable. Havalais (2013) argues that such explorations might open up ‘new structures of collaborative, participatory research.’ To explore whether this is indeed the case, we need to study how digital methods are shaped by actors within specific organizational contexts and how these methods shape the social worlds that they are mobilized to understand. As Ruppert et al put it, we need to attend to ‘specific mobilisations which allow the digital to be rendered visible and hence effective in particular locations’ (2013, p.31). Enter action research.

Action-researching digital methods: methodological description and reflection

In this section, we describe and reflect on the action research methods we used to undertake our study. We worked with two city councils and one city-based museums group, both to explore how digital methods might help them to understand and engage their publics and to achieve our aim of examining specific uses of digital methods in particular locations. In naming our approach action research, we point to our intention to work collaboratively with our partners, to engage in both research/inquiry and action/intervention. As a number of action researchers have observed, at the core of this approach is an intimate relationship between scholarly inquiry and practical or political activity or intervention, ‘such that the focus of inquiry arises out of, and its results feed back in to, the activity concerned’ (Hammersley, 2002; see also Freire, 1970 and Reason and Bradbury, 2001). Yet central to much debate about action research is the question of whether it is possible to sustain both action and research in equal measure. Whilst some subscribe to the

Greek privileging of research, or *theoria*, over action, or *praxis* (for example Polsky 1971), Hammersley argues that the subordination of inquiry to action is more common. As evidence of this, he cites Reason and Bradbury's assertion that 'the primary purpose of action research is not to produce academic theories based on action; nor is it to produce theories about action; nor is it to produce theoretical or empirical knowledge that can be applied in action; it is to liberate the human body, mind and spirit in the search for a better, freer world' (Reason and Bradbury, 2001, p.2; see also Carr and Kemmis 1986). Both models – the subordination of action to research and the subordination of research to action – are legitimate, in Hammersley's view. Drawing from our experience and following Hammersley, we suggest that the subordination of research to action may be expected in action research and in digital methods, because action, or the production of results (whether social change or digital data) is the ultimate aim of both. We elaborate on this proposition below.

Prior to the project we discuss here and as highlighted in the introduction, some of us carried out interviews with our partners and other public sector organisations (five in total), to examine whether and how they used digital methods to understand the concerns of their publics. This formed part of a larger study of the impact of digital developments on public engagement (Authors 2012). We were interested in comparing perceptions of these methods to perceptions of other methods adopted by public sector organisations to engage publics, such as surveys, consultations and citizens' juries. Through these interviews, we identified that organisations were using tools like TweetDeck (an application for managing, organizing and tracking Twitter accounts) or websites like Museums Analytics (a platform which produces summaries of relevant social media activity for museums) to manage and monitor social media activity. Their focus on social data informed our decision to do the same. These tools were not used systematically and most organisations were keen to expand their use of digital methods, to do one or more of the following: identify key 'influencers' and networks with which to engage; manage and analyse enquiries and feedback from the public about services; monitor relevant publicity; identify and analyse public conversation and opinion on local issues; identify and engage the public in policy development; and measure public engagement.

Given our interviewees' expressed enthusiasm to experiment with more digital methods and their limited resources to do so, we invited them to collaborate with us to explore a wider range of tools and methods than they already used. Two city councils and one museums group accepted the invitation. Others (such as a museum group) declined, because they believed that their existing 'manual' analytics (such as looking at discussions on city forums and on their own Facebook and Twitter accounts) were sufficient. For them, analogue approaches were felt to be more appropriate than digital methods for their public engagement needs. This decision was an early indication that digital methods are not always what small-scale public sector organisations need.

Through consultation with our partners and with a social media analytics expert, we identified some free or affordable tools with which to experiment. These included: NodeXL (<http://nodexl.codeplex.com/>), a freely available social network analysis tool created by The Social Media Research Foundation; Gephi (<https://gephi.org/>), an open source tool for network visualisation; and DataSift (<http://datasift.com>), a commercial, online tool that harvests data from a variety of social media platforms at low cost (at the time of writing, 20 US cents per unit, which can include up to 2000 tweets). NodeXL can access a limited number of platforms, one at a time, so adding DataSift to the toolset

allowed us to generate multi-platform data sets which could be exported into the other applications. We aimed to explore whether these tools would enable our partners to identify significant yet hitherto unknown influencers within their target communities with whom to engage, which they identified as one possible use of digital methods. In addition, we included IssueCrawler (www.issuecrawler.net), a free tool produced by the Digital Methods Initiative, in our toolset. IssueCrawler identifies issue networks – that is, networks linked by interest in specific issues, rather than social networks. Although designed initially for academic usage, non-profit bodies like councils and museums can also use it. We also trialled two commercial social media insights tools, one at the request of one of our partners, and the other as a comparison: Meltwater Buzz (<http://www.meltwater.com/products/meltwater-buzz-social-media-marketing-software/>) and Brandwatch (<http://www.brandwatch.com/>). Together, these tools enabled us to carry out investigations that covered the major categories of digital methods: social network analysis, issue network analysis, exploratory content analysis and visualisation of the resultant data sets. The toolset was expanded as the project progressed, because the free tools were experienced as difficult to use, because one of our partners wanted us to experiment with a commercial tool produced by a company from which they already licensed a traditional media monitoring platform, and because during the course of the research, it became apparent that fewer digital methods were being used by our partners than we believed as a result of the interviews. This suggests that our interviewees wanted to create the impression that they were more engaged in digital methods than was actually the case, an indication perhaps of the power of big data rhetoric to persuade people that this is what they should be doing.

Our inclusion of commercial tools within the project toolset indicates a number of things about the action-imperative of both action research and digital methods. First, it points to some of the problems our partners encountered using NodeXL, Gephi and IssueCrawler because of what they saw as the complexity of these tools; , we say more about this later. More importantly, our move to include commercial tools demonstrates the difficulty we encountered balancing the problems of digital methods with their potential, or research with action. In agreeing to collaborate with us, our partners hoped to find some data. They used digital methods because they were motivated by a desire to get results. To enable this, we used commercial tools, which were more effective than free tools. This move constructed us as intermediaries between the tools, their developers and our partners, something we discuss in greater detail below. It simultaneously compromised our initial intention of experimenting with free technologies which enable people to access digital data despite limited economic means and moved us closer to the action and further from the research of action research.

Over six months, we worked with our contacts in the communications teams of our partner organisations to experiment with these tools. Meetings were more frequent with the museums group than the other two partners, primarily because of the greater availability of staff therein. Thus what started as a collaborative and participatory project became less so for two partners, who subsequently had less opportunity to develop digital methods skills than we had all hoped. To compensate, for each partner, we produced a report summarising what social data we had found about their organisation through our experiments. These reports were intended as indications of what is possible with digital methods, rather than as comprehensive accounts of available data. We also produced a ‘guide to tools,’ which we shared with

all partners, to enable them to continue experimenting after the end of the project. In its introduction, we stated that there is much that is not known about how tools work, that choices about how tools are made shape the data they produce, and that access conditions constantly change. This represented an attempt to take discussion of the social shaping of digital methods into the public domain and to highlight some of the problems with these methods. This one page was followed by 46 pages of instructions, diagrams, and ‘how-to’ text. Sharing this guide was intended as a social intervention, to enable public sector organisations to engage in digital methods and so to realise our interest in exploring their potential. Producing instructions on how to use the tools constructed us not simply as intermediaries between tools and partners, but as advocates for the methods and tools discussed. As Law et al (2011) point out, this is how methods work – they need advocates in order to be adopted. Our advocacy of the potential of digital methods overshadowed our attempt to create spaces in which to discuss problems – there was one page of ‘problems’ compared to 46 pages of ‘potential’. Whilst we did not quite ‘go native’ (Malinowski 1922), the interventionist intent of our action research eclipsed our research intent in this instance.

To further compensate for two partners’ limited engagement in the experimental phase of the research, we ran a training workshop at the end of our project, which offered partners a hands-on opportunity to experiment with tools. It was attended by 13 representatives from a range of departments in our partner organisations. In both the workshop and the guide, we introduced more freely available social data analysis tools than those listed above, in order to equip our partners with quick-and-easy methods, as we had identified how difficult it was for them to dedicate time and effort to using the tools in our toolset. These included, for example, Social Mention (<http://www.socialmention.com/>) and Topsy (<http://topsy.com/>), social media search and analysis tools which aggregate content from social media sites and provide data such as numbers of comments on a given topic, sentiment, reach, top users and sources. As an indication of the simplicity of these tools, the Social Mention interface is shown in Figure 1.

The data generated on our project is social data drawn from social media platforms and other areas of the web where users contribute content, such as city-based forums or comments sections on local newspaper websites. There are other methods and tools that we could have used to access and analyse such data, all with distinct strengths, limitations and affordances. We were led in our choices by our own knowledge, the advice of experts, and the requirements of our partners. Tools needed to be free or cheap given our partners’ resource constraints, and the kind of analysis they facilitated was also an important consideration, as most partners had informed us that they were already doing some analysis, and we therefore sought tools which would allow them to advance their existing practice. These criteria changed as the research progressed. Committed to the ‘action’ element of action research, wanting to ensure that our partners’ participation in our project and use of digital methods produced the results they expected and responding to their changing requirements by suggesting alternative tools, we found ourselves in the position of intermediaries and advocates for digital methods tools, sometimes prioritising their potential over their problems, acting and intervening to ensure they got results. In the next section, we discuss three points which serve as further examples of this and of some of the difficulties we encountered in using our chosen tools to meet our partners’ needs.

Digital Methods in Practice: Three Issues

What's in a keyword?

At the beginning of the project, we suggested that our partners should identify topics of current concern on which to concentrate, as we felt that this might be a focused way of experimenting with digital methods. The museums group chose to focus on a) a current photography exhibition about nature and b) a set of online learning resources about museum education. One council chose to focus on a) council budget cuts and b) their new health and well-being strategy. The other city council was interested in investigating social data about the Tour de France, which would visit the region in 2014. This partner's second chosen topic was a new market which was due to open up in the city centre. In many ways, the choice of topics was not of great importance, as our intention was experimental.

We asked our partners to identify 20 keywords for each topic, words that they imagined their constituencies might use when commenting on them and their chosen topics. These would serve as starting points for our action research. Keyword selection is an important component of digital methods: an image of the phenomenon under investigation is created through the keywords that are used to describe it and this image is reflected in the results generated. This initial image is often refined throughout the iterative search processes. On our project, many of the keywords suggested by our partners in relation to their chosen topics produced no search results, as these terms were not necessarily used by the wider public in online conversation. Searches in DataSift using the keyword 'telehealth' (in relation to health and wellbeing) returned no results, as did specific phrases such as 'fortnightly bin collections' in relation to council budget cuts. Other keywords, such as 'Moor', the name of the new city centre market, produced results that were too broad to be useful. To ensure that some data was found, the research team used keywords provided by partners as starting points for identifying other, more widely used terms. For example, in relation to the photography exhibition, we used the Flickr search API to find search terms used by Flickr members to tag images that were also tagged with the keywords supplied by our partner. This intervention suggests that digital methods are only useful if users have relevant knowledge, in this case about keyword selection. In our case, we intervened in a way not initially intended, utilizing our expertise to overcome the limitations in our partners' knowledge in order to ensure that some data was found and so enable our partners to reflect on its potential usefulness. We intervened so that our partners got what they wanted from their collaboration with us, and so that digital methods delivered what we all expected them to: data. This is another example of the research team acting as intermediaries, prioritizing intervention over enquiry, action over research.

The need for expertise

The above discussion demonstrates that expertise in keyword selection is central to digital methods, as well as showing how combinations of keywords shape resulting data. Indeed, digital methods require expertise of many kinds. Manovich argues that there are three ways of relating to big data: there are 'those

who create data (both consciously and by leaving digital footprints), those who have the means to collect it, and those who have expertise to analyze it' (2011, p.10). As Manovich sees it, it is only this latter group, the analysts, who need expertise. boyd and Crawford (2012) argue that who has big data expertise determines both who controls the process and the 'knowledge' about the social world that results, knowledge which, as Law and others suggest, shapes the social world. But access to big data expertise is uneven, and this leads to new digital divides (or the reproduction of old ones). Our research aimed to confront the danger of a digital data divide by experimenting with free tools, but the use of these tools did not circumvent the danger of an expertise-based digital divide. On the contrary, we found, it served only to highlight it. For in contrast to Manovich's suggestion that only analysts require expertise, it was apparent in our research that expertise is also required in order to generate (or 'collect') data, as highlighted in our discussion of keyword selection above. Expertise is also needed to use data collection tools like DataSift, not only to understand their interfaces, but also the fields of data held in records returned by the APIs of platforms like Twitter and Facebook, as shown in Figure 2. Of course, data analysis also requires expertise. For NodeXL, certain technical knowledge is needed just to find it, as it is an Excel template file, not a standalone programme. Once opened, understanding of the terminology used is needed – for example, users need to know that choosing 'levels to include' means selecting whether to build networks of followers only (level 1), of followers of followers (level 2), and so on (and it is helpful if they also know that level 1 searches rarely reveal anything interesting). Then, users need to know how to make sense of search results (Figure 3 shows an example), how to manipulate results so they are visualized in meaningful ways, or how to export results into other tools, like Gephi, to produce visualisations.

This need for expertise was readily acknowledged by representatives from our partner organisations. One said 'you need to know the software inside out. You need to understand how to get into the data using that software' (Web Manager, Council Customer Services Department). Even the commercial tools that we trialled were experienced as difficult to use by our workshop participants, perhaps somewhat surprisingly given efforts made to produce usable graphical interfaces, a sample of which is produced in Figure 4 below. However, they still require certain kinds of expertise, for example in writing Boolean searches, and so require investments of both time (to gain expertise) and money (to purchase a license).

It might seem obvious that expertise is needed to use sophisticated tools like NodeXL, but our point is that the complexity of such tools, which are public in the sense that they are freely available, limits their usability by public organisations, because restricted resources constrain their ability to invest time into developing the skills required to use them. This is almost always a feature of the public sector, and exacerbated in austere times. As we have shown, to address this barrier, the research team intervened in various ways. We changed keywords through processes of iterative search and we brought new tools into the project toolset in the hope that they would be more successful in finding data and be easier for our partners to use. In doing this, we enabled partners to find data, as more data was found with tools introduced later in the research. We intervened because of perceived barriers to tool use and because all of us, the research team and our partners, engaged in action research with digital methods with the expectation that

data would be produced (and with an interest in whether produced data would produce change). Through our interventions, we played a role in producing and sustaining in our partners a commitment to using digital methods.

Data non-abundance

One of the most widespread assumptions about digital data relates to its volume; allegedly, there is lots of it (see for example Anderson, 2008). The major social media platforms themselves affirm the abundance of social data: in May 2013, Facebook (2013) reported 4.75bn pieces of content shared daily and at the same time there were reportedly an average of 400 million tweets posted every day (Twitter, 2013). Such assertions also serve to produce expectations that digital methods will find data. However, there is great disparity between the data that is available in relation to large-scale global topics and that which relates to small-scale local topics of the kind that our partners were interested in exploring. It is not the case that vast quantities of data are always there to be analysed, and finding relevant data in this mass can be challenging.

When setting out to find data about the topics chosen by our partners, one of our first steps was to explore online sources manually, through keyword searches for relevant conversations and by looking for key platforms used to discuss chosen topics. We also carried out web searches using lists of known individuals and groups provided by our partners and compiled the URLs of websites on which they contribute opinion. Our methods of manual investigation allowed us to identify important platforms, such as local city forums and comments sections of local newspaper websites. This process was not time efficient, nor did it capture a large sample of content. But it served the important purpose of showing us that relevant content existed. In contrast, as we started to use automated tools to search for data on a larger scale, the platforms that we had identified did not feature in results, and limited relevant data was identified. Data shortage was sometimes because of lack of expertise in keyword selection, discussed above, but even when efforts were made to improve keywords, results improved only slightly and several topics still returned few or no results.

On the whole, being local organisations, our partners were interested in finding local conversations and local ‘influencers’ with whom to engage. However, very little social data contains accurate location information. For this to be available, a location aware device with location services turned on is required, and users need to have agreed to their location being shared. Alternatively, location data can be derived from social media platforms themselves but again, such information is not widely available (Graham et al, 2013). Thus utilising geographic filters to limit harvested data to that which is generated in target areas brought limited success, as it diminished an already small pool of data and excluded relevant contributors (such as local people writing comments on newspaper websites, forums, blogs and in most cases Facebook and Twitter too) who were not sharing their geographical location in any way.

What’s more, tools are designed to source and analyse data in different ways, and these choices shape the resulting data. DataSift, for example, does not search the local platforms where our partners might find relevant data, like the city-based forums or regional newspapers’ comments sections which we identified in our manual search. The tool user can sometimes make decisions about where to search, but the choice is

limited to the platforms which the tool developer considers to be relevant and has included in the tool. On DataSift, the user chooses which social media to pay to access, from a limited list of major platforms. Our ability to intervene in relation to data non-abundance was more limited than in the previous two examples. However, we pointed out the absence of city-based forums amongst the platforms searched by the commercial tools to our contact at one of the commercial tool companies, who subsequently added these forums to the tool. Once again, we acted in order to produce results. But even before this move, the commercial tools offered more success in isolating local contributions than free tools.. It is hard to make a definitive statement about why this is, because many of the tools in question are black-boxed, a concept long used in science and technology studies to refer to the complexity of the technologies which, like methods, shape and constitute our social world, yet about the internal workings of which there is little common understanding (Pinch, 1992). Comprehending what is inside the black box of digital methods tools is important because of the effects that they have, because of their social life. But public understanding of technological functionality is often limited, partly because of black-boxing. In the concluding section, we argue that opening up the black box of digital methods is necessary in order for them to be useful for the public good.

Conclusion: on acting, researching and reluctantly having an impact

In this paper we have shown how a small-scale, qualitative study of what is assumed to be large-scale, quantitative data and its analysis can help to develop understandings of the ‘specific mobilisations’ of digital methods ‘in particular locations’ (Ruppert et al, 2013). Through our action research, we applied qualitative sensibilities and methods (Baym, 2013) to enable us to explore how digital methods are shaped by actors within specific contexts and how these methods shape the social worlds that they aim to understand. We used these methods to examine how groups who do not have the economic means to pay for commercial data services and who want to use available (and in this case, social) data might access it. Because of our interest in addressing these different questions, in our study, we attempted to remain open to both the potential and problems of digital methods, exploring empirically the assertions of digital data critics and advocates alike. We discuss potential and problems in much more detail elsewhere (Authors 2015 forthcoming). Here, we focus on how our combination of digital methods and action research worked together to limit our success in balancing our commitment to both potential and problems. We have pointed to moments when our commitment to ensuring our partners were able to access data meant that we sometimes became the advocates that Law et al (2011) argue that methods need, advocating for the digital methods about which we wanted to think critically. By acting to provide solutions to the difficulties encountered – improving keywords and introducing new tools – we produced (or at least perpetuated) in our partners a desire to engage with digital methods. Following Hammersley (2002), we suggest that, whilst different combinations of researchers, tools, partners and contexts might produce different results, the tendency to privilege action over inquiry is not unusual in action research, given its commitment to social change. We suggest that a tension between exploring potential and recognizing problems is also integral to working critically with digital methods, to efforts to simultaneously use and critique them. The use of digital methods is often motivated by a desire to produce results, so digital methods themselves produce the expectation that data will be found, that results will be produced, and that actions might be taken.

Acting to realise these expectations can sometimes get in the way of thinking critically about them.

Our partners tell us that the reports we shared with them, the guide to tools we created and their attendance at our end-of-project workshop have fed into the drafting of their social media policies. In addition, in one partner organisation, the communications team held workshops to introduce team staff to a range of digital methods. In this respect, the knowledge of digital methods which the project enabled has been useful, according to our key contacts, and the action we have taken has indeed produced results. We appear to have overcome a potential digital divide for a moment, in a very modest way. But we are ambivalent about having this impact. This is not only because of the double social life of digital methods, but also because it confirms that our modest efforts to open up a space to reflect critically with our partners on the social shaping of digital methods were overshadowed by more pressing concerns within the short time frame of our project, such as finding some data and doing some analysis. The challenge for the future, then, is to work on this tension between use and critique and to try to find ways to embed critical reflection more effectively in the action research process, the difficulties in doing so notwithstanding. A starting point for this might be public discussion of the double social life of digital methods, beyond the academy. We need to find ways to make it possible for resource-stretched public bodies and other community and public groups who are considering using these methods to discuss available tools and their limitations. Such discussion could serve as a space in which to consider the ways that digital methods could serve the public good.

Public uses of digital methods may have participatory consequences, as Havalais suggests, or they may have more troubling outcomes, because technologies for analysing digital data, especially those that are commercially developed, often enable regimes of governance and control (Trottier 2012), privacy invasion (Nissenbaum 2010) and transparency evasion (Turow 2012) (Baym, 2013; Ruppert 2013). For the purposes of our study, we conflated proprietary and free and open-source tools the workings of which are less black-boxed. We conclude by suggesting that, in order to consider whether digital methods are appropriate mechanisms with which to overcome emergent digital divides and whether they could be participatory, we may need to disaggregate not only tools, but also their methodological implementation and the different socio-technical environments in which they are mobilised. Whether public, community and social organisations can be digitally enabled through digital methods in the same way that other digital technologies have formed a part of digital enablement may be dependent on this disaggregation, as it may be that the biases written into the commercial digital methods tools used to perpetuate an increasingly knowing capitalism (Thrift, 2005) are absent from more public and open tools like NodeXL, Gephi and IssueCrawler. As digital methods become more widely enacted, these important questions about the differences amongst them, and their distinct normative and political effects, need further consideration.

References

- Anderson, C. (2008). The End of Theory: will the data deluge make the scientific method obsolete? *Edge*. Retrieved from:
http://www.edge.org/3rd_culture/anderson08/anderson08_index.html.
- Authors (2015 forthcoming).
- Authors (2012).
- Barnes, T. (2013). Big data, little history. *Dialogues in Human Geography*. 3(3), 297-302.
- Bassett, C. (2015 forthcoming). Plenty as a Response to Austerity?: Big data expertise, cultures and communities. *European Journal of Cultural Studies*.
- Baym, N. (2013). Data Not Seen: the uses and shortcomings of social media metrics. *First Monday*. Retrieved from:
<http://firstmonday.org/ojs/index.php/fm/article/view/4873/3752>.
- Boellstorff, T. (2013). Making big data, in theory. *First Monday*. Retrieved from:
<http://firstmonday.org/ojs/index.php/fm/article/view/4869>.
- boyd, d. & Crawford, K. (2012). Critical questions for big data: provocations for a cultural, technological and scholarly phenomenon. *Information, Communication and Society*, 15(5), 662-679.
- Carr, W. and Kemmis, S. (1986). *Becoming Critical*. Lewis: Falmer.
- Cote, M. (2014). Data Motility: the materiality of big social data. *Cultural Studies Review*, 20(1). Retrieved from:
<http://epress.lib.uts.edu.au/journals/index.php/csrj/article/view/3832>.
- Facebook (2013). Facebook's Facebook Page. Retrieved from:
<https://www.facebook.com/facebook>.
- Freire, P. (1970). *Pedagogy Of The Oppressed*. New York: Continuum.
- Gitelman, L. & Jackson, V. (2013). Introduction. In L. Gitelman (Ed.), *Raw Data is an Oxymoron*. Cambridge, MA: MIT Press.
- Graham, M., Hale, S. A., & Gaffney, D. (2013). Where in the World are You? Geolocation and Language Identification in Twitter. *Professional Geographer*.
- Hammersley, M. (2002). Action Research: a contradiction in terms? *British Educational Research Association*, Exeter, England, 12-14 September 2002. Retrieved from: <http://www.leeds.ac.uk/educol/documents/00002130.htm>.
- Havalais, A. (2013). Home made big data? Challenges and opportunities for participatory social research. *First Monday*. Retrieved from:
<http://firstmonday.org/ojs/index.php/fm/article/view/4876>
- Helles, R. & Bruhn Jensen, K. (2013). Introduction to Making data – big data and beyond'. *First Monday*. Retrieved from
<http://firstmonday.org/ojs/index.php/fm/issue/view/404>.
- Law, J., & Urry, J. (2004). Enacting the Social. *Economy and Society* 33(3), 390–410.
- Law, J. Ruppert, E. & Savage, M. (2011). The Double Social Life of Methods. CRESC Working Paper 95.
- Lowndes, V., & Squires, S. (2012). Cuts, collaboration and creativity. *Public Money & Management* 32, 401–408.
- Malinowski, B. (1922). *Argonauts of the Western Pacific: an account of native enterprise and adventure in the archipelagoes of Melanesian New Guinea*. London and New York: Routledge.
- Manovich, L. (2011). Trending: the promises and the challenges of big social data. Retrieved from:
http://www.manovich.net/DOCS/Manovich_trending_paper.pdf.

- Nissenbaum, H. (2010). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, Stanford University Press.
- Pinch, T. (1992). Opening Black Boxes: Science, Technology and Society. *Social Studies of Science*, 22(3), 487-510.
- Polsky, N. (1971). *Hustlers, Beats, and Others*. Piscataway: Aldine.
- Reason, P. & Bradbury, H. (2001). 'Introduction', in P. Reason and H. Bradbury (Eds.) *Handbook of Action research: participative inquiry and practice*. London, Sage.
- Rogers, R. (2013) *Digital Methods*. Cambridge, Massachusetts, The MIT Press.
- Ruppert, E. (2013). Rethinking Empirical Social Sciences. *Dialogues in Human Geography* 3(3), 268-273.
- Ruppert, E., Law, J. & Savage, M. (2013). Reassembling Social Science Methods: the challenge of digital devices. *Theory, Culture and Society*, 30(4), 22-46.
- Savage, M & Burrows, R (2009) Some Further Reflections on the Coming Crisis of Empirical Sociology, *Sociology*, 43, 762-772.
- Thelwall, M., Buckley, K. and Paltoglou, G. (2011) 'Sentiment in Twitter events', *Journal of The American Society for Information Science and Technology*, 62(2): 406-418.
- Thrift, N. (2005). *Knowing Capitalism*. London, Sage.
- Trottier, D. (2012). *Social Media as Surveillance: Rethinking Visibility in a Converging World*. Farnham: Ashgate Press.
- Turow, J. (2012). *The Daily You: How the New Advertising Industry is Defining Your Identity and Your Worth*. New Haven and London: Yale University Press.
- Twitter (2013). Celebrating #Twitter7. 21 March 2013. Retrieved from: <https://blog.twitter.com/2013/celebrating-twitter7>.
- Williamson, B. (2014). The death of the theorist and the emergence of data and algorithms in digital social research. *Impact of Social Sciences*, LSE Blog. Retrieved from: <http://blogs.lse.ac.uk/impactofsocialsciences/2014/02/10/the-death-of-the-theorist-in-digital-social-research/>.

Figure Captions

Figure 1. Social Mention interface

Figure 2. a table of results in DataSift, which requires knowledge of data fields to comprehend

Figure 3. The results of a search as displayed in NodeXL

Figure 4. A sample interface from one of the commercial tools