The redistribution of methods: on intervention in digital social research, broadly conceived

Noortje Marres

Abstract: This paper contributes to debates about the implications of digital technology for social research by proposing the concept of the redistribution of methods. In the context of digitization, I argue, social research becomes *noticeably* a distributed accomplishment: online platforms, users, devices and informational practices actively contribute to the performance of digital social research. This also applies more specifically to social research methods, and this paper explores the phenomenon in relation to two specific digital methods, online network and textual analysis, arguing that sociological research stands much to gain from engaging with it, both normatively and analytically speaking. I distinguish four predominant views on the redistribution of digital social methods: methods-as-usual, big methods, virtual methods and digital methods. Taking up this last notion, I propose that a redistributive understanding of social research opens up a new approach to the *re-mediation* of social methods in digital environments. I develop this argument through a discussion of two particular online research platforms: the Issue Crawler, a web-based platform for hyperlink analysis, and the Co-Word Machine, an online tool of textual analysis currently under development. Both these tools re-mediate existing social methods, and both, I argue, involve the attempt to render specific *methodology critiques* effective in the online realm, namely critiques of the authority effects implicit in citation analysis. As such, these methods offer ways for social research to intervene critically in digital social research, and more specifically, in redistributions of social methods currently ongoing in online environments.

Keywords: digital social research, social studies of science and technology, digital devices, online network analysis, online textual analysis, digital social methods

Introduction

As sociologists like to point out, the implications of technology for social life tend to be imagined in either highly optimistic or deeply pessimistic ways (Woolgar, 2002). Current debates about the implications of digitization for social research are no exception to this rule. The question of how digital devices,

The Sociological Review, 60:S1, pp. 139–165 (2012), DOI: 10.1111/j.1467-954X.2012.02121.x

and their proliferation across social life, transform social research is generating much interest today, and, as a consequence, the question of the 'social implications of technology' is now very much posed in relation to social research itself (Back, 2010; Savage *et al.*, 2010; boyd and Crawford, 2011). As it turns out, these discussions are no less susceptible to the polarizing effects of technology on the imagination, than, say, popular debates about the implications of cloning or robotics on society. While some propose that new technologies are opening up a golden age of social research, others argue that digitization has engendered a crisis for social research, creating a situation in which we risk to lose 'the human element' from view.

Both the optimistic and the pessimistic vision of digital social research start from a similar observation: digital technologies have enabled a broad range of new practices involving the recording, analysis and visualization of social life (Fielding *et al.*, 2008). Millions of blogs document everyday life on an ongoing basis; online platforms for social networking such as Facebook generate masses of data for social analysis; and applications of 'digital analytics' make it possible for everyone with access to these tools to analyse 'social behaviour' in real time. For the optimists, this situation implies a renaissance of social research: the new technologies and practices greatly enhance the empirical and analytic capacities of social research, and they render social research newly relevant to social life (Latour *et al.*, 2012). For the pessimists, the new digital sources of social intelligence announce not so much a rejuvenation of social research, but rather pose a serious threat to established traditions and forms of sociological research (Savage and Burrows, 2007). From this vantage point, the proliferation across social life of new technologies for recording, analysing and visualizing social life masks an underlying trend of a very different nature. These technologies are leading to the privatization of social research: they enable the displacement of social research to the corporate laboratories of big IT firms.

In this paper, I would like to unsettle this opposition between a utopian and dystopian imagination of digital technology in social research. I would like to contribute to debates about the implications of digitization for social research by exploring the concept of the *redistribution of research*. This notion has been put forward in the social studies of science and technology (STS) to complicate our understanding of the relations between science, technology, and society (Latour, 1988; Rheinberger, 1997; see also Whatmore, 2009). It highlights how scientific research tends to involve contributions from a *broad* range of actors: researchers, research subjects, funders, providers of research materials, infrastructure builders, interested amateurs, and so on. Scientific research, according to this notion, must be understood as a *shared accomplishment* of a diverse set of actors. This idea has clear implications for digital social research: it suggests that it may be a mistake to try and locate digital social research in a single domain, be it 'the university', or 'everyday practices like blogging', or 'the private laboratories of large IT firms'. Instead, we should examine how, in the context of digitization, the roles of social research are being distributed between a range of different actors: between researchers, research subjects, digital tech-

nologies, and so on. Moreover, the concept of redistribution directs attention to a possible implication of digitization for social research: digitization may be unsettling established divisions of labour in social research. If we use blogs in social research, does this mean that we are partly delegating the task of data collection to bloggers?

Here I would like to focus here on the redistribution of a specific element in social research, namely methods. Digitization has special implications for the role and status of social research methods in particular (Fielding *et al.*, 2008; Rogers, 2010; Adkins and Lury, 2009). Views on this matter, too, diverge: some propose that digital technology inaugurates an age of methodological innovation, as new technologies for data collection, analysis and visualization enable the further elaboration of existing methods and the development of new ones. Others are more inclined to emphasize the 'return of the same' masked by such claims to newness, proposing that the 'new' digital methods continue along the same path as the 'quantitative revolution' of the 1960s and 70s (boyd and Crawford, 2011; Uprichard *et al.*, 2008). These observations are no less pertinent than the optimistic and pessimistic diagnoses flagged above, but on the issue of method too, there seems to be potential in side-stepping the 'false choice' between an utopian and a dystopian diagnosis, and to examine instead whether and how digitization enables new ways of *distributing* methods among different agents involved in social research. Social methods, too, may be understood as a shared accomplishment, involving contributions of researchers, research subjects, technologies, and so on (Rogers, 2009). The question is how the digital inflects this circumstance.

The issue of the redistribution of methods is a slippery one, as the contributions of different agents to the enactment of methods are hard to pin down: to return to the above example, why would we call blogs agents of data collection, rather than data points in our source set? On what grounds? To prevent being paralysed by general questions like this, I will explore the redistribution of method here in a contextual and empirical way, namely by examining two online platforms for social research: Issue Crawler, a web-based application for network analysis which has been online for 10 years now, and a tool of online textual analysis that is currently under development, provisionally called The Co-Word Machine. Both of these tools adapt social research methods to the online environment, namely network and textual analysis, and more precisely, co-citation and co-word analysis.[1] And both of these platforms can be said to undertake a 'redistribution' of social research methods: they transpose onto the Web methods that have long been championed in social research and, in doing so, they come to rely on a different set of entities in the enactment of this method, such as Web crawlers and online data feeds. The translation of methods of network and textual analysis into online environments, I will emphasize, enables a form of critical intervention in digital social research: to implement these methods online is to offer a distinctive variation on more prevalent applications of methods of network and textual analysis in digital networked media. The overall aim, then, is to get a more precise sense of the space of intervention

opened up by digital social methods – *of method as intervention* – online. First, however, I would like to revisit in more detail the current debate about the implications of digitization for social research.

The digitization of social life and the redistribution of social research

The ongoing debate about the implications of digital technology for social research has directed attention to three significant features of digitization. No doubt the most important one is the *proliferation of new devices, genres and formats for the documentation of social life*. The last decade has seen an explosion of digital technologies that enable people to report and comment upon social life, from photo-sharing via Flickr to the public gossip of Twitter. Such online platforms allow users to publicize their accounts of everyday life like never before, in the form of simple text or snapshots taken with mobile phones. Especially interesting about the new devices from a sociological perspective is that they enable *the routine generation of data about social life as part of social life* (Fielding *et al*., 2008; see on this point also Marres, 2011). 'Social media' platforms, that is, embed the process of social data generation in everyday practices, whether in the form of people 'live' commenting on an event via Twitter to the smart electricity meters that record fluctuations in domestic energy use. Finally, the two previous developments cannot really be understood without considering the development of online platforms and tools for the *analysis* of digital social data.

These days, most online platforms come with 'analytics' attached: a set of tools and services facilitating the analysis of the data generated by said platforms, from blog posts to Facebook friends. In this respect, what is especially significant for social research about online platforms for 'user-generated content' is that they actively support the adaptation of these platforms for purposes of social research. An example here is Yahoo Clues, a recently launched online platform that makes data generated by the Yahoo search engine available for analysis, allowing 'you to instantly discover what's popular to a select group of searchers – by age or gender – over the past day, week or even over the past year'.[2] Providing access to a searchable database of search engine queries, Yahoo Clues makes available for analysis an arguably new type of social data, in the form of millions of queries that people perform as part of everyday life. And as Yahoo Clues allows its users to break down popular queries in terms of searcher profiles (gender, age, geographic location), it enables a distinctively social form of analysis. It also provides an example of the 'relocation' of social research enabled by digitization, as it formats social analysis as a popular practice that 'anyone' might like to engage in.

Social theorists have been hard pressed to provide an integrated assessment of these various developments and their implications for social research. Some authors have sought to affirm the new popular appeal of social research, suggesting that we are today witnessing a radical expansion in the range of actors,

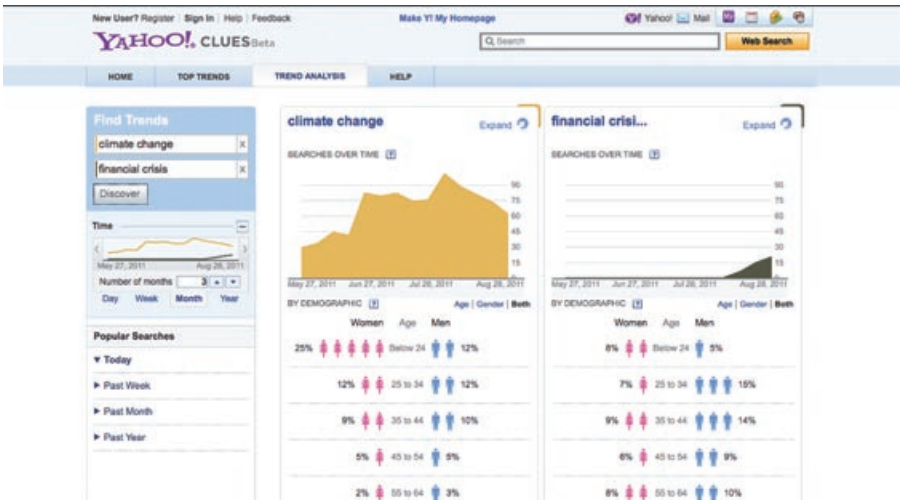


Figure 1: *Yahoo Clues: a new beta service that gives you a unique 'behind the scenes' look into popular trends across the millions of people who use Yahoo! to search each day (July 2011).*

devices and settings caught up in the recording, reporting and analysis of social life. Some sociologists have been tempted to see in social media platforms a clear case of 'non-professional researchers enthusiastically engaging in the recording and reporting of social life' (my formulation). This would suggest that digitization is occasioning a revival of amateur-led social research, invoking memories of the English Mass Observation Movement, with its armies of lay people who documented scenes of everyday life in notebooks and questionnaires during the 1930s and 40s (Hubble, 2006; Savage, 2010). But others – indeed, in some cases the same authors – are more drawn to the dark side of this vision. Thus, Savage and Burrows (2007), in their influential article on 'The Coming Crisis of Empirical Sociology', prophesized that digitization signals the demise of sociology as a public form of knowledge. In their account digitization, in spite of popular appearances, enables the *concentration* of social research capacity in a few well-resourced research centres, most notably of big IT firms. In this view, the wide popularity of online platforms for the collection, annotation and analysis of social data makes possible the concentration of research in a few hubs of the digital economy, equipped for the central storage, processing and valuation of these data.

As has often been pointed out, the optimistic and the pessimistic diagnosis of a social phenomenon, while in some ways strictly opposed to another, may in other ways be neatly aligned (Haraway, 1991; Woolgar, 2002). As we know from the social study of consumer culture, dynamics of popularization and infrastructural concentration are by no means anti-thetical. As Celia Lury (1996, 2004) observed, popular fashion brands like Nike are marked by prolif-

eration and unification, by the combination of an open-ended multiplicity of Nike-inflected social practices and a centralized orchestration of the phenomenon. To observe, then, that the spread of digital devices for the recording and analysis social life occurs simultaneously with the concentration of control over the infrastructure that enables it is to note an all too familiar feature of post-industrial societies. It is just that, in the context of digitization, these dynamics are proving increasingly relevant to social research itself. But here I would like to argue that by concentrating on this overarching issue of the *displacement* of research capacity – to society at large, or the IT industry – we risk losing from view another, more fine-grained dynamic: the *redistribution* of social research *between* actors involved in social research. Rather than rushing to decide which sector of society will prove to be the biggest 'winner' – which will strengthen its position the most as a consequence of the digitization of social research? – we must then consider a more open-ended and complex process, namely that of the reconfiguration of the relations between a diverse set of agents caught up in social research.

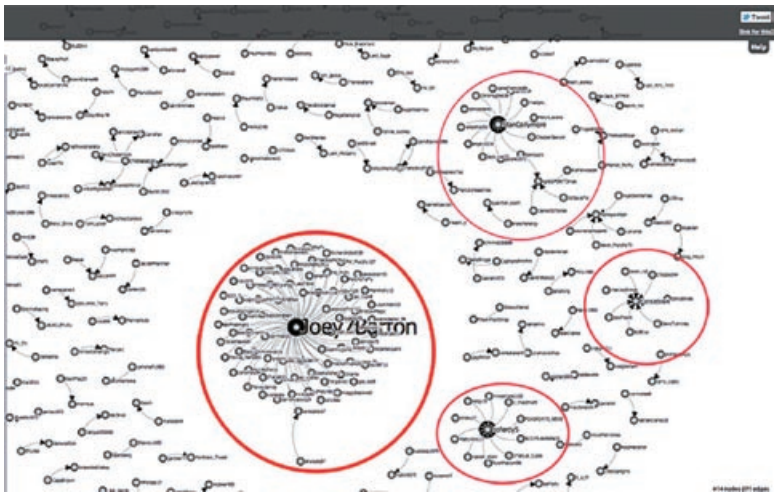
The notion of the 'redistribution' of research has been put forward in STS and related fields to highlight processes of exchange *between* actors involved in social research. The notion emphasizes that the production of new knowledge and new technologies tend to involve complex interactions and transactions between a whole range of actors inside as well as outside the university, including research subjects, funding bodies, technological infrastructures, researchers, and so on. Research and innovation, then, is also a matter of the *transfer* of information, materials, and also more complex things like 'agency', between the various actors involved in research: when subjects agree to be interviewed or offer samples, when an institution allows a researcher into its archive, certain transactions occur that are critical to the production of new knowledge or technology. One example here is focus group research: this form of research relies on contributions from a range of actors, from research subjects, to research subject recruitment agencies and focus group moderators (Lezaun, 2007). Rather than assume that focus group research is conducted either 'in the university' or 'in the corporate sector', it seems to make more sense to consider how this methodology enlists actors from different practices and domains, from marketing to government, activist organizations and academic research, and enables transactions among them. Indeed, social studies of focus group research have shown that the invention of the focus group in 1940s America enabled social research to take on new roles in society, among others as advisers on civic opinion (Lezaun, 2007; Grandclement and Graglio, 2010). It also involved the development of new 'infrastructures' of social research, such as focus group research centres.

The concept of the 'redistribution of social research' has a number of implications for the debate about the consequences of digitization for social research. It suggests that some of the assumptions informing the question of the displacement of research capacity, from the university to society, or from the public university to private industries, may be too simplistic. Quite possibly the idea

of the self-sufficient academy has *always* been a myth (Latour, 1988; Button, 1991; Callon *et al*., 2009). For a long time already, academics have not been the only or even the main protagonists of research, as other actors have historically played active roles in the production of knowledge (Latour, 1988; Law, 2004). It is just that the conventional understanding of science and innovation makes it difficult to acknowledge the active contributions of 'non-scientists' as meaningful contributions to research and innovation, without problematizing the status of our knowledge. Going against this understanding, the concept of the redistribution of social research proposes to define social research as inherently a collective undertaking, involving a diverse set of actors in a variety of roles. Processes of inquiry, from this vantage point, are best understood as *inherently* distributed among a whole range of agencies, involving active contributions from research subjects, the experimental apparatus, funders of research, and so on (Latour, 1988; Rheinberger, 1997; Law, 2009).

From the standpoint of an understanding of social research as distributed, the question of *displacement* of research capacity – away from academia; towards popular culture or industry – is not the most relevant question to ask. Rather than trying to decide in what *singular* location research capacity is today most advantageously located, we should examine what digitization means for the distribution of roles in social research *between* various actors in and outside the university. Especially important about digitization, from this vantage point, is that it may well be unsettling divisions of labour in social research. Emerging practices of online social research seeking to take advantage of the new social data made available by platforms like Facebook and Twitter provide a case in point. Digital sociology student Sam Martin, for instance, turned to Twitter to analyse the racial abuse row over the prosecution of England footballer John Terry.[3] Using various applications from Google Docs to Yahoo Pipes and the Twitter API ('application programming interface'), Martin culled messages mentioning John Terry from Twitter over a four-day period in February 2012. Using a programme called 'TagExplorer' she produced a network map of 'top-conversationalists', which notably included 'Queens Park Ranger Captain and Footballer' Joey Barton, who was present at the pitch when the racial abuse incident occurred (see Figure 2).

This type of online research, which adapts social media applications to the purposes of social research, can be said to redistribute social research in various ways. Most notably, arguably, is its reliance on the social media platform Twitter itself: Twitter ranks tweets and tweeters according to the number of followers, tweets, and re-tweets, and in visualizing the corpus of messages using the measure of 'topconversationalists', Martin's small study arguably replicates some of the measures that are implicit in the medium under scrutiny. We should also note the various research tools and applications that allowed her to extract tweets from Twitter and visualize them, like Tagexplorer: these instruments, as well as the 'developer community' from which they sprang, here come to play a notable role in the organization of social research, and so did, arguably, the army of tweeters who in this study got a say on framing phenomena as



Colour online, B&W in print

Figure 2: *Top conversationalists, the John Terry debate on Twitter, visualization using TAGSExplorer, 3–7 February 2012, by Sam Martin.*

sociologically significant (by following some Twitter contributors rather than others).

Digital social research may then entail particular redistributions of sociological research. Taking up digital online tools, sociological research is likely to enter into working relations with platforms, tool developers and analytic and visual devices which are operating in contexts and developed for purposes that are not necessarily those of sociology (Marres and Weltevrede, forthcoming). In examining such redistributions in digital social research, we can ask a question about the implications of digitization for social research that is at once more specific and open-ended than the question about displacement: to what extent does digitization enable *renegotiations of divisions of labour* in social research between the various actors involved? At issue, then, is not only which institution or sector gets to define what social research is, and to occupy the 'top spot', but rather what relations between a range of different actors is enabled by particular, emerging digital social research practices. The notion of the redistribution of social research, furthermore, directs attention to *a much broader* set of actors and entities caught up in the process of the digitization of social research, including but not limited to: online platforms, users, databases, design agencies, algorithms, IT companies, digital culture commentators, information formats, social movements, and so on (see on this point also Madsen, 2012). The division of labour between users, devices and researchers in the conduct of social research, we then say, is being unsettled, contested and redefined in complex but quite specific ways.

The idea of the redistribution of social research can provide some useful conceptual guidance, I think, in examining the implications of digitization of

social research. It differs from the thesis of the 'displacement' of social research, highlighted above, in at least four ways.

First, to consider the redistribution of social research is to shift attention from the *external* relations of social research to its *internal* relations. The displacement diagnosis posits a fairly strict separation between academic social science and its various outsides – industry, social life, the public. To argue that research capacities are moving away from academia to somewhere else is to accentuate the distinction between academic and other forms of social research. By contrast, a redistribution perspective is concerned with the division of roles between actors inside and outside the university in the production of social research (Adkins and Lury, 2009; Savage *et al.*, 2010).[4] It entails a relatively loose definition of social research, to which various skills and competencies may contribute. Secondly, a redistributive understanding of social research implies a shift in perspective from *ready-made sociology* to *sociology in-the-making*. The digitization of social research, we could say, renders newly relevant a classic insight of the social studies of science and technology: our analysis of knowledge production changes radically as soon as we shift our attention from the status of social research *as a finished product*, to ongoing processes of social research (Latour, 1988).

To conjure up the spectre of the 'corporatization' or 'popularization' or 'democratization' of social research, is to build an argument that derives its normative force from a focus on outcomes. By contrast, if we focus on divisions of labour in digital social research, we explore rather how digitization may affect and inform the *conduct* of social research, and the normative charge of our exploration here derives from the extent to which these processes are still to a degree undecided, contested, multiple. Thirdly, and relatedly, the notion of redistribution leads us to question the distinction *between the conditions or 'context' of social research and its content.* Debates about the consequences of digitization of social research often concentrate on changes that affect the 'material base' for social research, that is, the technologies and forms of data storage on which it relies. However, of many of the features of digital social research it is actually quite hard to say whether they affect only the conditions or the substance of social research or both or neither: does Twitter research primarily signify a change of conditions in social research, as tweets can be extracted from Twitter so much faster and in quantities that are so much larger than used to be the case in popular discourse analysis (boyd and Crawford, 2011; Leavitt, 2009)? Or does the very meaning of the concept of social discourse change now that we mean by it the broadcasting of one-liners by active individuals in 'real-time' (Niederer and Van Dijck, 2010)?

Fourthly and finally, a focus on redistribution rather than displacement has implications for how we understand our own role as social researchers. That is, the practical or normative roles that we are able to envision for social research, or what we might call their 'scope of intervention', is very different depending on which of the two perspectives we adopt, displacement or redistribution. From a redistributive perspective, the principal question becomes how we may

most relevantly intervene in shifting distributions of social research capacity. Here, the main point is not to paint big canvas total pictures of the unlikely future we desire for social research and the likely one that we must fend off. Rather, the question becomes where and how, given the type of redistributions of social research that are currently ongoing, we can most pertinently add a different ingredient that might change the wider mix of social research. A focus on social research *methods* appears to be especially productive in this regard.

The redistribution of social research methods: five views

Method is an important mediator of divisions of labour in social research, and this is no less the case in digital social research. The devising of new research methods, of course, has long been a strategy of choice for those attempting to establish privilege, or claim precedence or newness in science, and digital social research is no exception to this either. As in other fields, debates in social research about methodology have long served as a key site and proxy for much more comprehensive controversies about the future direction of the field, with much of the 20th-century methodology contests having been dominated by the pitching of quantitative versus qualitative sociology, with the Positivismusstreit between Karl Popper and Jurgen Habermas as an illustrious example. Methods, then, offer a means to conjure up and establish particular versions of social research, and this in turn tends to involve the attempt to enforce particular divisions of labour in social research. Qualitative social research, for instance, proposes to grant much more initiative to research subjects, while much quantitative research endeavours to create a greater role for standardized tools of data collection, such as the survey, as a way to guarantee the commensurability of data.

In the area of digital social research, methods are invoked to such effects as well.[5] There have been some audacious claims about the opportunities for methodological innovation enabled by online networked media, such as the ability to detect patterns in user activity on the Web which may indicate or predict real-time events, like an onslaught of the flu (Rogers, 2009; Mohebbi *et al.*, 2011). And in this context, too, qualitative and quantities methods are pitched against one another, as claims are made back and forth about the relative advantages of, for instance, digital ethnography versus large-scale online survey research (boyd and Crawford, 2011). The Internet has also been said to favour particular social methods over others, such as unobtrusive or non-interventionist methods like content analysis (Lee, 2000; Carslon and Anderson, 2007). Here I cannot do justice to these various methodology debates, but discussions about digital social research methods provide an especially useful prism through which to approach the issue of the redistribution of social research. Different views on the implications of digitization for social methods imply very different understandings of what redistributions of research capacity are possible in this context, both empirically and normatively speaking. They therefore

provide a useful way to identify different options in this regard. In this section, I will present some different views on the digitization of the particular methods of networks and textual analysis, so as to set the stage for further discussion of one approach which I think has special affordances for intervention in digital social research.

It is possible to order different views on the implications of digitization for social research methods along a spectrum, which starts on one end with a minimal redistribution of research capacity and moves to a maximum redistribution on the other end. The left side of this spectrum is marked by a conservative position that is sceptical about the possibility that social methods are undergoing any significant transformation in digital environments, let alone something like a 'redistribution of methods'. This position, which might be dubbed 'methods-as-usual' can be recognized in an argument recently put forward by the eminent Chicago sociologist Andrew Abbott, who proposed that for anyone who is well versed in social research methods, the newness of the new, online media is very much overstated.[6] Abbott emphasizes that the social methodologies incorporated into digital devices like search engines, most notably network and textual analysis, are pretty standard fare, at least for trained sociologists, and has called the search engine Google 'basically a concordance machine', which matches key-words (queries) to target contexts, and which relies on 'rather routine' additional measures of network analysis, such as in-centrality, to determine the authority of sources, something which has little new to offer to sociologists who have long been familiar with such measures. This view focuses specifically on the formal metrics built into digital devices, and does not consider how these metrics are adapted to or informed by other features of digital devices, as for instance the use of 'live' data or feedback mechanisms. Indeed, it does not really consider the possibility that social research methods may be transformed by virtue of their insertion in a digital networked environment. One could accordingly say that, from this perspective, only one redistribution of research capacity has occurred, in that popular online devices now have social research methods built into them. But on the whole no real redistribution of *methods* takes place: social research methods themselves are not really affected by their uptake in digital online media.

A second view differs significantly from this, and is associated with the new network science informed by mathematics, physics and computing science. This body of work is principally concerned with the opportunities that online media offer for further development of large-scale network and textual analysis, and may accordingly be called 'big methods'. It proposes that digitization has made possible new developments in the *modelling* of networks and textual worlds, and this in large part because of the very large data-sets that digital media technologies make available. The vast databases that have been built over the last decade by search engine companies, gaming industries, Internet service providers and social media platforms create opportunities to significantly expand the analytical and empirical power of network science. They enable the further development of what Duncan Watts and others (Newman *et al.*, 2007) refer to as 'the

analysis of real-world network dynamics' (see also Lazer *et al*., 2009). Contrary to methods-as-usual, this methodological programme can be said to undertake a redistribution of methods of sorts. This new network science namely favours a new set of techniques for data collection and analysis, which entail a different division of labour between research subjects, data collection devices, and analysts in social research. To put it somewhat crudely, the approach seeks to maximize the role of mathematical techniques, at the expense of research subjects. Thus, in their introduction to the New Network Science, Newman and Watts argue that the social data generated by digital platforms are 'more amenable to the kinds of techniques with which physicists and mathematicians are familiar', and offer a welcome substitute for survey data, and other all too 'social' forms of data (Newman *et al*., 2007).

The new network science reinstates a classic opposition of social research, that between subjective and objective data. Like many others, Newman *et al.* (2007) locate the opportunities that digitization offers for social research in the *type* of data that now become available for social analysis: namely transactional data, which 'record the activities and interactions of the subjects directly' and are thus routinely generated as part of social activities by digital devices, from loyalty cards to search engines (see on this point also Latour, 1998; Rogers, 2009; Savage and Burrows, 2007). Newman *et al.* (2007) give a classic positivist justification for relying on this type of data, arguing that they are much more objective and, as such, offer a welcome substitute for the 'subjective' data generated by surveys, making it possible to avoid reliance on the active contributions of erratic human subjects to data collection.[7] In their account, then, data provided by research subjects are not quite reliable data, something which in their view challenges the validity of network analysis as a whole: 'the respondent data are so contaminated by diverse interpretations of the survey instrument, along with variable recollection or even laziness, that any inferences about the corresponding social network must be regarded with scepticism' (Newman *et al.*, 2007: L-6). The rise of social media like email, blogs and Facebook here makes possible the *rejection* of user-generated data for purposes of social research, and a redistribution of research capacity towards online registrational devices.

A third and fourth approach are respectively called 'virtual methods' and 'digital methods', and they can be distinguished from the former two in that they are explicitly concerned with the changing relations between social research, its devices and objects in digital online environments. These two approaches offer, however, very different accounts of these changes. The 'virtual methods' programme, developed by Christine Hine (2000, 2005) and others in the early 2000s, focused on the opportunities opened up by the transposition of qualitative social research methods into digital online environments. Its main concern was the digital transformation of *our own sociological methods*, that is, the way methods like discourse analysis and ethnography were and could be transformed by their application in a new context. In focusing mostly on the fate of qualitative methods, Hine's approach to virtual methods makes the opposite manoeuvre from the new network science: it seeks to maximize the role of

interpretative subjects in social research, defining the experience of this subject as one of the principal empirical objects of virtual social research. As Hine (2002) puts it: 'ethnographers of the Internet cannot hope to understand the practices of *all* users, but through their own practices they can develop an understanding of what it is to be *a* user' (2002: 54). More generally speaking, the virtual methods approach is concerned with the *digitization* of social research methods, that is, with the translation of methodologies that sociologists define as their own into online environments (Rogers, 2010). This is to recognize a significant but limited redistribution of methods: here, the role of new entities, like web users, in the performance of social method is very much acknowledged, as everyday Internet users are seen to do things online that are similar to fieldwork (taking notes, documenting practice, checking out a strange, new social world). However, such redistributions of social method are here only explored insofar as they occur in the realm opened up by the sociologists' research itself: researcher, research subjects, mediating infrastructures, tools used, and so on.

In adopting this strategy, virtual methods do not address the wider issue of the general uptake of social methods in digital online environments, and the consequences of this for the shape and outlook of digital social research. It is this issue that the digital methods programme formulated by Richard Rogers and others (Rogers, 2009) explicitly takes up. This approach proposes the dominant digital devices, search engines chief among them, can be adapted for purposes of social research, and accords to these devices the capacity to inform the development of new methods of social research. Because of their large, dynamic data sets, sophisticated algorithms and feedback possibilities, search engines, Rogers argues, are able to devise forms of social analysis that were not possible before, which he terms 'natively digital' (see also Weltevrede, n.d.). Digital methods, then, propose that social research should take advantage of the analytic and empirical capacities that are 'embedded in online media'. These can be adapted to purposes of social research, by developing online research tools that run on top of web devices, like Google. The Googlescraper, for instance, adapts Google to conduct work frequency analysis in source sets delineated by the user.[8] This methodological programme of *repurposing* entails a particular redistribution of social research methods, namely towards *devices*: in proposing to adapt existing online devices for purposes of social research, their capacities of data collection, analysis and feedback, come to be incorporated into social and cultural research. As the Digital Methods Initiative proposes to import dominant online tools for data collection, analysis and visualization into social research – or at least parts thereof – devices that constitute the *context* of digital culture come to actively inform the content of social and cultural research.

Arguably, the Digital Methods Initiative more than any other approach discussed above seeks to come to terms with the redistribution of methods in digital environments. Recently, sociologists have recognized that online environments foster a range of tools and practices that qualify as instruments as social research, acknowledging that methods lead a 'social life' online (Savage *et al*., 2010). But the Digital Methods Initiative proposes an empirical programme that deliber-

ately deploys this circumstance, seeking to render it analytically useful for social research. However, in its above formulation, this approach nevertheless could be said to share a blind spot with the first two approaches that I discussed above. Just as with the methods-as-usual perspective and the 'big methods' of the new network science, digital methods can be seen to bracket the issue of the *re-mediation* (Bolter and Grusin, 2000) of social methods in digital online media. As mentioned, Rogers defines the methods enabled by online digital devices as 'natively digital', proposing that they have no clear correlate in the pre- or non-digital world. In making this claim, the DMI programme statement does not really consider, or even downplays, the question of how the uptake of existing social research methods in digital environments entails a *refashioning* of these methods.[9] This question, however, seems to me all too relevant if we are to appreciate the type of *interventions* that social research may become capable of in the broader context of the redistribution of social methods in online environments.

The notion of the 're-mediation of methods' is useful, I want to propose here, in that it directs attention to the ways in which prevailing digital devices have methods built into them in which we can recognize those of social research. The foundational article in which Google founders Larry Page and Sergey Brin outlined the central idea behind the new search algorithm, Pagerank, does not only cite a famous sociologist of science, Robert Merton, but it also makes an informed critique of the limitations of sociological forms of network analysis, or as the case may be, citation analysis (Page *et al.*, 1999). Below I will further discuss the particular re-mediation of citation analysis undertaken by Google. Attending to such re-mediations of specific social methods in the digital context, I want to propose, opens up a particular mode of intervention for social research itself. Insofar as predominant digital devices apply existing social methods, this may render newly relevant existing sociological *critiques* of these methods. The re-mediation of social methods in the digital context, then, opens up a space of critical intervention for engaged social research in the broad context of the online application of social methods. In the remainder of the piece, I will discuss the methodological strategies that involved the development of two digital research tools along these very lines. The Issue Crawler, and an online application of co-word analysis currently under development, the Co-word machine. By considering how two devices re-mediate social methods, we get an idea of the *digital forms of methodology critique* they enable.

Issue Crawler: from co-citation to co-link analysis

Issue Crawler is an online platform for the location, analysis and visualization of hyperlink networks on the Web. Launched in the early 2000s, Issue Crawler was intended to enable the location and analysis of 'issue networks' on the Web, as it uses hyperlink analysis to delineate sets of pages dealing with a common theme that are connected by hyperlinks (Marres and Rogers, 2008). But the tool

has since been used in a variety of projects of online network analysis, including organizational networks (allowing organizations to answer questions such as 'how central are we in this area?') as well as the longitudinal study of online networks, as in the analysis of the rise of Obama and his social media campaign sites in the US democratic election network of 2008 (Borra, 2008; see Figure 3). Using the campaign sites of all democratic presidential candidates as starting points, this last study used IssueCrawler to conduct a series of scheduled crawls, which plotted the emergence of a highly ordered network on the Web, with Obama social media campaign sites dominating the entire network towards the end of the election period. Such network dynamics are arguably Web specific, insofar as the reconfiguration of material network relations can be analysed in real time. But the method on which Issue Crawler relies to demarcate hyperlink networks is based on a classic form of network analysis: co-citation analysis. As an implementation of this specific social research method, the design of Issue Crawler is clearly informed by the context in which the platform was developed.[10]

In the late 1990s, as I already mentioned above, the rise of the Internet was widely interpreted as an opportunity to apply methods of citation analysis in the new medium, and to adapt this classic method for the analysis of hyperlink structures (Scharnhorst and Wouters, 2006). In this period, the rise of Google and its famous Pagerank algorithm, which relies on in-link measures to rank sources in its query return lists, made newly relevant debates of citation analysis

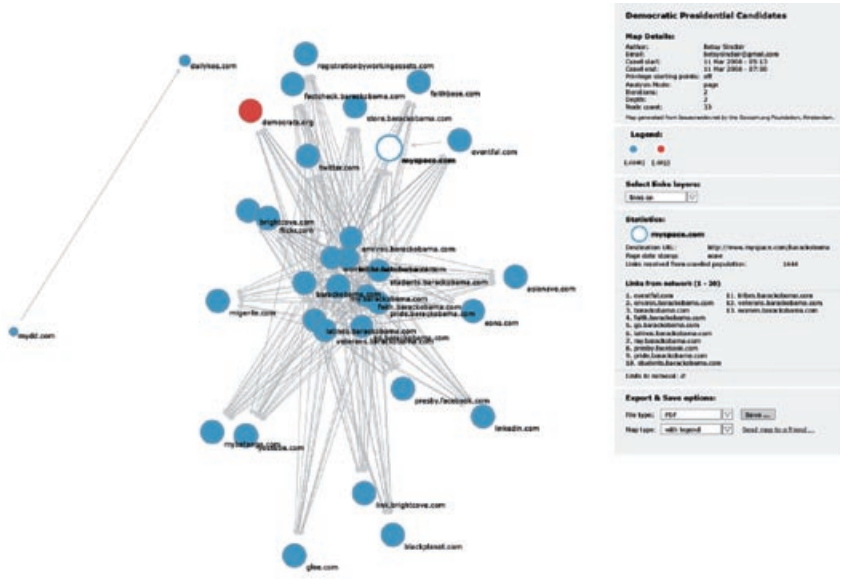


Figure 3: *Issue Crawler Map: The rise of Obama and Web 2.0 in the Democratic Presidential Candidates network, Besty SinClair, March 2008*

Colour online, B&W in print

that had been developed by sociologists of science from the 1960s onwards. Larry Page's foundational article makes a specific argument about the re-mediation of citation analysis enabled by the Web, which in his view makes it possible to address *a shortcoming of this method:*

> There has been a great deal of work on academic citation analysis. Goffmann has published an interesting theory of how information flow in a scientific community is an epidemic process. [. . .] But the reason Pagerank is interesting is that there are many cases where simply citation counting does not correspond to our common sense notion of importance. For example, if a web page has a link of the Yahoo home page, it may be just one link but it is a very important one. This page should be ranked higher. Pagerank is an attempt to see how good an approximation to 'importance' can be obtained just from the link structure. (Page *et al.*, 1998)

Arguably, this issue of 'source authority' had already been discussed in citation analysis (MacRoberts and MacRoberts, 1988), and accordingly the degree to which Google's brand of hyperlink analysis contains an actual innovation can be debated. However, as methods of citation analysis were being re-invented as methods of hyperlink analysis, the question was raised whether and how *critiques* of citation analysis transferred into the online environment. This was – in one of those stories one can tell about tools and methods – the question that Issue Crawler was made to address.

In the 1960s and 70s, sociologists had voiced concerns about citation analysis that now proved all too relevant to the methodological innovation proposed by Google. Citation analysis, it had been argued back then, enables a potentially perverse authority dynamic, according to which well-cited sources get cited more simply because they are already well-cited (Small, 1973; see also Callon *et al.*, 1983).[11] Any large number of citations tends to generate more of them, these now classic critiques proposed, resulting in a situation in which sources are considered authoritative simply by virtue of their authority, and accordingly processes of the valuation of knowledge are captured by social dynamics of popularity, and risk to become divested from more substantive processes of valuation. This classic critique touched on issues of clear relevance to the new search engine algorithms, like Google's (Marres and Rogers, 2000): these algorithms, too, attributed authority to sources on the basis of the level of recognition implied by the overall number of hyperlinks they received, independent from content. In developing the methodology of Issue Crawler we then drew on this specific methodological critique of 'the authority of authority', in order to develop an alternative approach to hyperlink analysis, one that draws on co-citation analysis (Marres and Rogers, 2000).

In some respects, then, Issue Crawler simply transposed an old methodological solution into a new context. Co-citation analysis was developed in the 1960s as an alternative to the standard citation measure of the overall number of citations received. Rather than seeking to determine the overall authority of individual sources, co-citation analysis seeks to delineate clusters of relevant sources by identifying sources that are jointly linked to by other sources. Applying this

method to hyperlinks, Issue Crawler sought to introduce a substantive measure of relevance into hyperlink analysis. Issue Crawler deploys the method of co-link analysis in order to undercut the authority effects to which citation and network analysis are vulnerable: instead of assigning value to the overall number of links that sources receive, co-link analysis seeks to locate 'topical clusters' of sources, by identifying co-links in a thematic neighbourhood, or as we called them 'issue networks'. As is clear from the example in Figure 3, Issue Crawler has not necessarily been successful in foregrounding dynamics of relevance at the expense of dynamics of authority. Arguably, indeed, the more insightful issue networks located with Issue Crawler include a clear element of authority, though this is not always the case (Marres and Rogers, 2008 discuss some exceptions).

However, it is also important to note that in transposing co-citation analysis onto the Web, Issue Crawler transformed this classic method in some important respects. Before the Web, co-citation analysis was by its very nature limited to the analysis of scientific data-bases, most notably the Science Citation Index. Even as this method sought to challenge authority dynamics, it inevitably rendered itself dependent on institutional demarcations of the relevant fields, in this case scientific fields. For this reason, co-citation could *not* include all the sources to which citations directed it: the scope of its analysis was limited to the sets of sources included in official scientific databases. The Web, by contrast, presents us with *networks* of databases, and as such, it opens up the possibility of analysing a much broader array of sources in real time, generating data-sets that are much more *heterogeneous* than those of citation analysis (Marres and Rogers, 2000; Muniesa and Tchalakov, 2009).[12]

In using co-link analysis to locate thematic networks on the Web, Issue Crawler does not only transpose a particular method into the online environment, but also a specific *methodology critique*. In advocating co-citation analysis, sociologists did not only seek to address a problem with methods of citation analysis in themselves, or with questionable citation behaviours, whereby sources mainly recognize already authoritative sources, thus aggravating the popularity effect. In the pre-digital context, critics of citation analysis specifically targeted the ways in which citation analysis *amplified* these popularity effects: their concern was that science policy would increasingly rely on these methods, as research councils took up citation measures, in their attempt to render their modes of assessment more evidence-based (Leydersdorff, 1998). Similarly, the issue with search engines is not just that, in applying measures of in-link centrality, they help to generate more authority for already authoritative sources.[13] At issue is a whole complex of behaviours: by privileging sources with a high overall in-link count, search engines encourage linking behaviours that consolidate authority dynamics, and the modification of user trajectories to a similar effect (Introna and Nissenbaum, 2000; Vaidhyanathan, 2011). In networked environments, then, it is especially obvious that *multiple* agencies have a part to play in the enactment of 'social methods'.[14] To put it differently, in the digital context social methods must clearly be defined as a *distributed* accom-

plishment, and our attempts to intervene critically in this context must be informed by this circumstance.

The co-word machine: from co-word analysis to online issue profiling

In questioning the dominance of authority dynamics on the Web, back in the late 1990s, and the role of devices like Google in enabling this, however, I clearly had little idea of what we were up against. In retrospect it can seem naive to expect that a methodology developed by a minoritarian movement in the sociology of science, like co-citation analysis, could be rendered effective in digital networked spaces, which were just then emerging as key hubs of the global information economy. Indeed, recent developments in this area, most notably the rise of social media platforms like Facebook and Twitter, can be taken as evidence that the medium has gone 'the other way'. Reputational dynamics, whereby things become more widely liked by virtue of being liked, have become very much the currency of online media (Onnela and Reed-Tsochas, 2010; Gerlitz and Helmond, 2012). The social network, in which actor-alliances are formed largely independent from content – and *not* the 'issue-network', with its topical dynamics of the thematic clustering of sources – has become the key organizational form associated with the Internet (for the distinction between issue- and actor-network, see Marres and Rogers, 2008). However, social media platforms also highlight the limits of our earlier argument in another, less ironic sense: social media have proven that networks driven by reputational logics are very well capable of organizing content, in ways that do *not* necessarily reproduce 'the tyranny of reputation'. The rise of these platforms has been accompanied by the proliferation of tools of content analysis and visualization. Figure 4, for example, provides a word frequency analysis of action terms on Facebook, showing the relative prominence of such terms in a selection of Facebook groups.

Social media, then, have proven to be no less adaptable to the purposes of content analysis than social network analysis. Nevertheless, I think that our initial intuition still holds: online digital environments are in need of alternative measures that can provide a counter-weight to dominant popularity metrics. On closer inspection, many current instruments of online content analysis, like tag clouding, have not really attenuated authority effects. They tend to rely on versions of the 'overall citation count' too: they bring into view what (or who) is most mentioned, followed, liked and so on, in a given data set at a given moment. Tag clouds, and other online applications of textual analysis and visualization perpetuate the preoccupation with the most cited or most popular, and these instruments can thus be said to reproduce the authority effect in another form. After the rise of social media, the question then remains how to develop alternatives to reputational measures: the question is still that of how to move beyond 'purely social' mechanics of authority, popularity or celebrity, and get to more substantive dynamics of relevance. But in this context, too, existing sociological critiques of research methods may offer a useful resource: debates

Figure 4: *Tag cloud analysis, Facebook is for joiners*
Source: Lonneke van der Velden and Clare Lee, Project Facebook, DMI Summerschool, July 2010 (https://wiki.digitalmethods.net/Dmi/TrainingProgramProjectFacebook).

about the majoritarian bias in textual analysis, and the development of alternative forms of 'discourse analysis' have been ongoing in sociology for several decades. Here I would like to single out one such alternative method, namely *co-word analysis*, as this method was explicitly developed by sociologists of science and technology developed to enrich citation analysis and possibly by extension hyperlink analysis.

Co-word analysis was devised in the 1980s by the actor-network theorist Michel Callon and others as a way to expand the project of co-citation analysis. It was developed to locate 'pockets of innovation' in science, using textual analysis to locate especially active thematic clusters of sources in the scientific literature.[15] Co-word analysis did this by measuring the rise and fall of keywords, and the associations among them, in a corpus of scientific articles (Callon *et al*., 1983; Whittaker, 1989). Relying on the keywords used to index the articles in scientific databases to build a lexicon, co-word analysis offered a way to determine which were the most 'active' key words, and word associations in the corpus. It provided a way to measure which keywords and keyword associations *varied* significantly in their mentioning and relations over a given period. In trying to determine the most 'happening' themes, this method was expressly designed to locate 'buzz' or 'live content' in the scientific literature, *but it did this without relying on popularity dynamics*. Indeed, terms that were mentioned with a constantly high frequency were automatically deleted from the set of active terms: the key indicator was not frequency of mentioning but *variation* in mentioning (and association) (Callon *et al*., 1983).

In recent years, the method of co-word analysis has been transposed onto the Web, with various online applications deploying the methodology to visualize word associations in online data sets, such as corpi of email messages or twitter messages (Danowksi, 2009; www.infomous.com). In the online context, co-word analysis promises to offer an alternative to word frequency analysis, the method of which it seems fair to say spread like wildfire, also into the social

sciences, on the back of tag clouding tools.[16] Co-word analysis determines the relevance of terms by measuring the strength and intensity of relations among them: only words that appear frequently *and* that appear together make it onto co-word maps. Co-word analysis, as mentioned, tries to purge its analysis of terms that are merely popular: it excludes terms that appear frequently but in random association with others. For this reason, co-word analysis promises to offer an alternative to the majoritarian logics of word frequency, which make a term appear larger and more visible the more often it appears. This method may help us move beyond the popularity ranking or hit list, that most visible testimony to the tyranny of reputation, as is evidenced by the alternative visual format proposed by Callon and colleagues for co-word analysis (see Figure 5). Crucially, moreover, online co-word analysis does away with popularity without sacrificing *liveness,* or rather liveliness. Co-word analysis, too, aims to deliver the most *happening* content (see also Marres and Weltevrede, forthcoming). But it does this by deploying an alternative measure: not the safety of a large number of mentioning, but fluctuations in the presence of words and word associations is key.

Together with colleagues in Amsterdam, we are now working to develop a Co-Word Machine that deploys co-word analysis for the online location and visualization of 'issue language'. In transposing co-word analysis into the online context, however, a number of issues arise which may either weaken or strengthen the analytical and critical capacities of this method, depending on how we deal with them, and how they will continue to play out. First of all, online environments offer an opportunity which adherents of co-word analysis could only theorize in the 1980s. For Callon and his colleagues, the chief attraction of co-word analysis was its promise to help advance 'our search for the trans-disciplinary, trans-institutional problematic networks that we want to identify' (Callon *et al*., 1983: 196) However, in the 1980s co-word analysts were frustrated

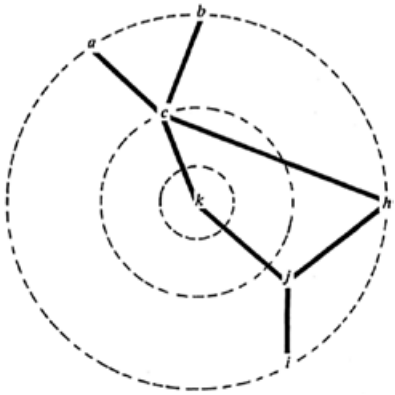


Figure 5: *Co-word visualization*
Source: Callon et al. (1993).

in this project by the limits of the databases and genres to which they applied their method. As in the case of co-citation analysis, co-word analysis relied on scientific databases, and because the genre of the scientific article was so different from those current in other fields (the policy report, the newspaper article, the petition), there was no reliable way to track terms *across* discursive spheres. Online networked media provide a great opportunity to address this limitation, as one distinctive feature of these media is precisely the significant genre contamination across fields (which organization does not have a blog?). In this environment, co-word analysis, too, may be applied to far more *heterogeneous* data-sets (Marres and Rogers, 2000).

However, the Web also poses serious challenges for co-word analysis, among others because of the widely divergent ways of indexing content prevalent in the medium. In this respect at least, classic co-word analysis had it easy, as it could rely on professional indices – keywords used by institutions like the Science Citation Index to index scientific articles – to locate emergent vocabularies. In online media, most applications rely on self-indexing – on keywords, or tags, provided by users marking up self-generated content. This inevitably raises issues of reliability and comparability, and in this respect, digital tagging practices drive home a basic but important point made by the American journalist Walter Lippmann (1997 [1922]) in his classic analysis of newspapers: any factual report is only as good as the sources from which it derives its findings (such as the National Office of Statistics). In this respect, co-word analysis certainly is *not* free of the problems associated with digital devices like tag clouding, which, as the name says, tend to rely on tags used to mark up online content, by bloggers and other users. In the case of co-word analysis as well, our results will only be as good as the classificatory practices on which we rely. We are returning then, to the issue of the distributed accomplishment of digital methods: online textual analysis builds on the contributions of a whole host of agents, from the availability of tagging features, to the taggers who actually mark up online content, the analytical instruments used to analyse these tags, visualization modules, and so on. In order to intervene relevantly in online social research, broadly conceived, I have argued here, we should recognize such assemblages of users, devices and informational practices, as the relevant unit of 'methodological innovation' in social research.

Conclusion

In online environments, the distributed nature of social research is especially hard to deny. User behaviours, information formats and digital devices that are embedded in the medium are likely to leave an imprint on social analysis. Something that applies to other research practices too is then rendered explicit in online social research: here, social research is *noticeably* marked by informational practices and devices not of its own making, from the analytic measures built into online platforms (eg numbers of links, number of mentionings,

follower counts), to the visual forms embedded in visualization modules (the tag cloud). Online social research is then visibly a distributed accomplishment. This circumstance, I have argued, does not only pose problems for social research but also offers opportunities for the development of social research methods. Digitization enables a broadening of the agencies playing an active role in the enactment of social methods, broadly conceived: in this context, a wide range of actors including platform users and analytic devices like search engines come to play a part in the collection, analysis and presentation of social data. And this redistribution of methods in digital social research opens up a space of intervention for social research.

Social methods, I have argued, are a key instrument through which wider divisions of roles in social research are being curated in online environments. Prominent digital devices like Google and Twitter and Facebook, and the users and developers enrolled by these platforms, today actively inform the enactment of social methods online. The types of data platforms make available, the measures and formats on which they rely in communicating this data (rankings, follower counts and clouds), and the wider informational practices in which they are taken up (Facebook members visualizing the network of their Facebook friends): all of these elements inform the performance of 'social methods' in digital networked environments. The contours of these 'methodological spaces online' are not necessarily easy to determine, as platform settings change, and users changing their allegiance to a new device. However, these assembled devices, settings and actors open up a particular space of intervention in digital social research: if specific digital social methods are a distributed accomplishment – such as the 'overall citation count' that is materialized in Google and other platforms – then sociology may seek to intervene in the relations among entities that sustain these methods, by proposing alternative methods and distributions thereof. Web-based applications of co-link analysis and co-word analysis, the Issue Crawler and the Co-word machine currently under development, aim to do just this. In so doing, they extend some of the long-standing normative projects of sociological research into digital environments, such as the commitment to methods that privilege substantive dynamics of relevance over purely social or reputational ones, or what we could call 'post-social methods'.

As noted, there is a strong tradition in social research of seeking to bracket the effects of the methods deployed by 'the social actors themselves': many social researchers have become experts in devising tricks that make it possible to ignore the active contribution of research subjects (and objects) to the organization of data and the framing of methods. But digital networked environments provide opportunities to explore different possible approaches to the distributed nature of social research and its methods. As online social research forces us to acknowledge the contributions of digital devices, practices and subjects, to the enactment of social research, it can be taken as an invitation to move beyond 'proprietary' concepts of methods, that is, beyond the entrenched use of method as a way to monopolize the representation of a given field or aspect of social

reality. A redistributive approach to social research redefines methods as involving the combination and coordination of diverse competencies: classification, visual design, automated analysis, and so on. Behind debates about the unreliability of data generated by research subjects, and the 'mess' of self-indexed online content, there lies a debate about the redistribution of methods between researchers, devices, information and users, in online environments. Which is also to say, the debate about the digitization of social methods is perhaps most productively approached as a debate about *participatory* research methods.

Acknowledgements

I would like to thank Les Back, Erik Borra, Sabine Niederer, Nirmal Puwar, Richard Rogers, Emma Uprichard and Esther Weltevrede for their comments on earlier versions of this article.

Notes

1 Both of these methods have been central to the development of actor-network theory and in focusing on the re-mediation of these methods, I am also exploring how online research tools translate methods of actor-network theory into networked digital media. On doing so, I will join others in arguing that digitization offers opportunities for a generalization of this sociological research programme (Latour, 1998; Law, 2008; Savage, 2010; Latour *et al.*, 2012).

2 'New Yahoo! Clues Launches', posted 29 June 2011, http://www.ysearchblog.com/2011/06/29/new-yahoo-clues-launches/

3 See http://twitterabused.com/2012/02/09/visualising-twitter-networks-john-terry-captaincy-controversy/

4 The notion of the redistribution of social research in the digital context is both inspired by and deviates from the idea of the double social life of methods proposed by Savage *et al.* (2010). Whereas the latter proposes that social research methods are both deployed in social science and in society at large – as for instance 'the survey' – the idea of the redistribution of research directs our attention to shifting relations *between* agencies inside and outside the university.

5 One redistributive issue requires special attention: digital social research entails a reshuffling of roles between human and technical elements, and as such it raises the question of which delegation of roles to new actors or devices are exactly occurring, what their significance is, and what their implications for the analytic and empirical capacities of social and cultural research (Niederer and van Dijck, 2010; see also Bach and Stark, 2005).

6 Andrew Abbott, 'Googles of the Past: Do Keywords Really Matter?', lecture, Department of Sociology, Goldsmiths, 15 March 2011.

7 They note: 'For most of the past fifty years, the collection of network data has been confined to the field of social network analysis, in which data have to be collected through survey instruments that not only are onerous to administer, but also suffer from the inaccurate or subjective responses of subjects. People, it turns out, are not good at remembering who their friends are, and the definition of a 'friend' is often quite ambiguous in the first place' (Newman *et al.*, 2007: L-5).

8 https://tools.issuecrawler.net/beta/scrapeGoogle/

9 This notion of re-mediation was put forward by Bolter and Grusin (2000) in an effort to shift the debate about digital culture beyond yes/no exchange which pitched two sterile positions

against one another: either new media merely offered old culture in a new jacket, or it enabled the invention of radically new forms of culture. Rejecting both positions, Bolter *and* Grusin proposed to focus instead on how older cultural forms underwent a process of *refashioning* in new media. I am proposing here that this notion can be usefully adapted to make sense of the digital social research methods.

10 Issue Crawler was developed between 1999 and 2002 by the govcom.org foundation in Amsterdam, which is directed by Richard Rogers and of which the author was a founding member. www.govcom.org.

11 This dynamic is in some ways similar to a classic sociological problematic, discussed by Tocqueville, of 'the tyranny of reputation'. According to this wider dynamic, ideas gain influence for the reason of being well regarded, a circular dynamic in which substantive considerations of the ideas in question do not necessarily enter.

12 In some sense, online hyperlink analysis enabled a move beyond the database. In this respect, the technique of crawling the Web allows for a renewed engagement with a classic sociological concern of actor-network theory: the issue of the pre-ordering of data, as what prevents sociology from engaging with heterogeneous ontologies.

13 Issue Crawler also engages with issues which in retrospect we can designate as issues of public sociology: its methodology concentrates on a publically accessible metric, hyperlinks, and its archive of all located networks, dating back to 2001, is available to all users.

14 Issue Crawler also seeks to put this situation to positive effect. The quality of its network maps depends on the knowledge implied in the hyperlinks that it analyses: Issue Crawler can only provide us with 'telling networks', if sources in the network link intelligently, ie if they identify issue-protoganists and alliances among them by way of hyperlinks.

15 More specifically, co-word analysis was developed as a way to deal with the problem that co-link analysis reproduced a reputational logic in spite of itself. This problematics is all too relevant in relation to Issue Crawler: this platform too can be said to reproduce popularity and authority effects, for various reasons: because of its demarcationist approach, because of hyperlinking reproducing authority effects, and because platform users want to know 'who is the most popular source'. In this respect, the reproduction of reputational dynamics by Issue Crawler is itself partly a social effect, ie it is a consequence of the distributed nature of digital social research: the effect can partly be traced back to 'reputational linkers', and the research agendas of the users of Issue Crawler.

16 There are a number of related tools for visualizing word frequency analysis, like the Dorling visualization, and one of my favourites, the Bubble line.

References

Adkins, L. and Lury, C., (2009), 'Introduction to special issue "What is the empirical?" ', *European Journal of Social Theory*, 12: 5–20.

Bach, J. and Stark, S., (2005), 'Recombinant technology and the new geography of association', in R. Latham and S. Sassen (eds), *Digital Formations: IT and the New Global Realm*, 37–53, Princeton, NJ: Princeton University Press.

Back, L., (2010), 'Broken devices and new opportunities': re-imagining the tools of qualitative research', NCRM Working Paper Series, 08/10.

Bolter, J. and Grusin, R., (2000), *Remediation: Understanding New Media*, Cambridge: MIT Press.

Borra, E., (2008), 'The web as anticipatory medium, blogpost', http://erikborra.net/blog/2008/12/the-web-as-an-anticipatory-medium/#more-709.

boyd, d. and Crawford, K., (2011), 'Six provocations for big data', paper presented at Oxford Internet Institute's *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society,* 1–17.

Button, G., (1991), 'Introduction', in G. Button (ed.), *Ethnomethodology and the Human Sciences: A Foundational Reconstruction*, 1–9, Cambridge: Cambridge University Press.

Callon, M., Courtial, J., Turner W. and Bauin, S., (1983), 'From translations to problematic networks: an introduction to co-word analysis', *Social Science Information*, 22: 191–235.

Callon, M., Lascoumes, P. and Barthe, Y., (2009 [2001]), *Acting in an Uncertain World: An Essay on Technical Democracy*, Cambridge: MIT Press.

Carlson, S. and Anderson, B., (2007), 'What are data? The many kinds of data and their implications for data re-use', *Journal of Computer-Mediated Communication*, 12 (2), http://jcmc.indiana.edu/vol12/issue2/carlson.html

Danowski, J., (2009), 'Network analysis of message content', in K. Krippendorff and K. Bock (eds), *The Content Analysis Reader*, 421–430, Thousand Oaks, CA: Sage Publications.

Fielding, N., Lee, R. and Blank, G., (2008), 'The Internet as research medium: an editorial introduction to *The Sage Handbook of Online Research Methods*', in N. Fielding, R. Lee and G. Blank (eds), *The Sage Handbook of Online Research Methods*, 3–20, London: Sage.

Gerlitz, C. and Helmond, A., (2012), 'The like economy: social buttons and the data-intensive Web', *New Media and Society*, forthcoming.

Grandclement, C. and Gaglio, G., (2010), 'Convoking the consumer in person: the focus group effect', in D. Zwick and J. Cayla (eds), *Inside Marketing: Practices, Ideologies, Devices*, Oxford and Cambridge, MA: Oxford University Press, http://dx.doi.org/10.1093/acprof:oso/9780199576746.003.0005

Haraway, D., (1991), 'Cybermanifesto: science, technology and socialist-feminism in the late 20th century', in *Cyborgs, Simians and Women: The Reinvention of Nature*, 149–182, New York: Routledge.

Hine, C., (2002), *Virtual Ethnography*, London: Sage.

Hine, C. (ed.), (2005), Virtual Methods: Issues in Social Research on the Internet, Oxford: Berg.

Hubble, N., (2006), *Mass Observation and Everyday Life: Culture, History, Theory*, Basingstoke: Palgrave Macmillan.

Introna, L. and Nissenbaum, H., (2000), 'Shaping the Web: why the politics of search engines matters', *The Information Society*, 16 (3): 1–17.

Latour, B., (1988), *The Pasteurization of France*, trans. A. Sheridan and J. Law, Cambridge, MA: Harvard University Press.

Latour, B., (1998), 'Thought experiments in social science: from the social contract to virtual society', *1st Virtual Society? Annual Public Lecture*, http://www.artefaktum.hu/it/Latour.htm

Latour, B., (2005), *Reassembling the Social: An Introduction to Actor-network-theory*, Oxford: Oxford University Press.

Latour, B., Jensen, P., Venturini, T. and Boullier, D., (2012), 'The whole is always smaller than its parts: a digital test of Gabriel Tarde's Monads', *The British Journal of Sociology*, forthcoming.

Law, J., (2004), *After Method: Mess in Social Science Research*, London and New York: Routledge.

Law, J., (2008), 'On STS and sociology', *The Sociological Review,* 56 (4): 623–649.

Law, J., (2009), 'Assembling the world by survey: performativity and politics', *Cultural Sociology,* 3 (2): 239–256.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D. and Van Alstyne, M., (2009), 'Computational social science', *Science*, 6 February, 5915 (323): 721–723.

Leavitt, A. (ed.), (2009), 'The Iranian election on Twitter: the first eighteen days', *Web Ecology Project*, http://www.webecologyproject.org/2009/06/iran-election-on-twitter/ (5 March 2011).

Lee, R. M., (2000), *Unobtrusive Methods in Social Research*, Buckingham: Open University Press.

Leydersdorff, L., (1998), 'Theories of citation?', *Scientometrics*, 43 (1): 5–25.

Lezaun, J., (2007), 'A market of opinions: The political epistemology of focus groups', *Sociological Review*, 55: 130–151.

Lippmann, W., (1997 [1922]), *Public Opinion*, New York: Free Press Paperbacks, Simon & Schuster.
Lury, C., (1996), *Consumer Culture*. Cambridge: Polity Press.
Lury, C., (2004), *Brands: The Logos of the Global Cultural Economy*, New York and London: Routledge.
MacRoberts, M. H. and MacRoberts, B. R., (1988), 'Problems of citation analysis: a critical review', *Journal of the American Society for Information Science*, 40 (5): 342–349.
Madsen, A. K., (2012), 'Web visions as controversy lenses', in A. Carusi, A. Sissel Hoel and T. Webmoor (eds), Special Issue on Computational Picturing, *Interdisciplinary Science Reviews*, 37 (1).
Marres, N., (2011), 'The cost of involvement: everyday carbon accounting and the materialization of participation', *Economy and Society*, 40 (4): 510–533.
Marres, N., (2012), 'The experiment in living', in C. Lury and N. Wakeford (eds), *Inventive Methods: The Happening of the Social*, 76–95, London: Routledge.
Marres, N. and R. Rogers (2000), 'Depluralising the Web and repluralising public debate: the case of the GM food debate on the Web', in R. Rogers (ed.), *Preferred Placement: Knowledge Politics on the Web*, 113–135, Maastricht: Jan van Eyck Editions.
Marres, N. and Rogers, R., (2008), 'Subsuming the ground: how local realities of the Ferghana Valley, the Narmada Dams, and the BTC pipeline are put to use on the Web', *Economy and Society*, 37 (2): 251–281.
Marres, N. and Weltevrede. E., (forthcoming), 'Scraping the social? Issues in real-time research', *Journal of Cultural Economy*.
Mohebbi, M., Vanderkam, D., Kodysh, J., Schonberger, R., Choi, H. and Kumar, S., (2011), 'Google Correlate Whitepaper', http://correlate.googlelabs.com/whitepaper.pdf
Muniesa, F. and Tchalakov, I., (2009), 'What do you think a simulation is, anyway? A topological approach to cultural dynamics', Working Paper, http://www.atacd.net/
Newman, M., Barabási, A. and Watts, D., (2007), *The Structure and Dynamics of Networks*, Princeton, NJ: Princeton University Press.
Niederer, S. and Van Dijck, J., (2010), 'Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system', *New Media and Society*, 12 (8): 1368–1387.
Onnela, J.-P. and Reed-Tsochas, F., (2010), 'Spontaneous emergence of social influence in online systems', *Proceedings of the National Academy of Sciences*, 107 (43): 18375–18380.
Page, L., Brin, S., Motwani, R., and Winograd, T., (1999), 'The PageRank Citation Ranking: bringing order to the Web. Technical Report', Stanford InfoLab. http://ilpubs.stanford.edu:8090/422/
Rheinberger, H.-J., (1997), *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*, Stanford, CA: Stanford University Press.
Rogers, R., (2009), *The End of the Virtual: Digital Methods*, Amsterdam: Amsterdam University Press.
Rogers, R., (2010), 'Internet research: the question of method', *Journal of Information Technology and Politics*, 7 (2/3): 241–260.
Savage, M., (2010), Identities and Social Change in Britain since 1940: The Politics of Method, Oxford: Oxford University Press.
Savage, M. and Burrows, R., (2007), 'The coming crisis of empirical sociology', *Sociology*, (41): 885–899.
Savage, M., Law, J. and Ruppert, E., (2010), 'The double social life of methods', CRESC Working Paper Series. No. 95, http://www.cresc.ac.uk/our-research/cross-theme-research/social-life-of-methods
Scharnhorst, A. and Wouters, P., (2006), 'Web indicators – a new generation of S&T indicators?' *Cybermetrics*, 10 (1).
Small, H., (1973), 'Co-citation in the scientific literature: a new measure of the relationship between two documents', *Journal of the American Society for Information Science*, 24 (4): 265–269.
Uprichard, E., Burrows, R. and Byrne, D., (2008), 'SPSS as an "inscription device": from causality to description?', *The Sociological Review*, 56 (4): 606–622.

Vaidhyanathan, S., (2011), *The Googlization of Everything (And Why We Should Worry)*, Berkeley, CA: University of California Press.

Weltevrede, E., (n.d.), 'Studying society, not Google: repurposing Google for social and cultural research', Department of Media Studies, University of Amsterdam, ms.

Whatmore, S. J., (2009), 'Mapping knowledge controversies: science, democracy and the redistribution of expertise', *Progress in Human Geography*, 33 (5): 587–598.

Whittaker, J., (1989), 'Creativity and conformity in science: titles, keywords, and co-word analysis,' *Social Science in Science*, 19: 473–496.

Woolgar, S., (2002), 'Introduction: five rules of virtuality', in S. Woolgar (ed.), *Virtual Society? Technology, Cyberbole, Reality*, Oxford: Oxford University Press.

